

**Facultad de Humanidades y Artes**

**Universidad Nacional de Rosario**

**Escuela de Posgrado**

**ANÁLISIS MORFOLÓGICO AUTOMÁTICO  
DE LA INTERLENGUA DE LOS APRENDIENTES  
DE ESPAÑOL COMO L2**

**Tesis presentada para optar por el título de Doctor en  
Humanidades y Artes con Mención en Lingüística**

**Carolina Tramallino**

**Directora: Zulema Solana**

**Agosto de 2016**

## TABLA DE CONTENIDOS

Página | 2

### INTRODUCCIÓN

Lingüística computacional.....	pág 6
Objetivos.....	pág 7
Fundamentación.....	pág 8
Estado de la cuestión.....	pág. 8
Hipótesis .....	pág 9
Etapas y Metodología de la investigación.....	pág. 10
Herramientas.....	pág. 11
Aportes.....	pág. 12
Organización.....	pág. 12

### CAPÍTULO 1: LA TEORÍA DE LA INTERLENGUA

1.1. Concepto de intelengua.....	pág. 15
1.2. Características de la interlengua.....	pág. 16
1.3.¿Cómo incide la lengua materna en la adquisición?.....	pág. 17
1.4. Clasificación de errores según su causa.....	pág. 19
1.4.2. Errores intralingüísticos.....	pág. 19

### CAPÍTULO2: ESTUDIOS DE ADQUISICIÓN DE SEGUNDAS LENGUAS

2.1. El análisis contrastivo.....	pág.21
2.2. El análisis de errores.....	pág. 23
2.3. Consideraciones finales.....	pág.25
Glosario de términos.....	pág.25

### **CAPÍTULO 3: EL CORPUS**

3.1. Organización del corpus.....	pág. 28
3.2. Los sujetos.....	pág. 29
3.2.1. ¿Por qué llamamos “segunda lengua”?.....	pág.31
3.3. ¿Por qué es importante el corpus?.....	pág.32
3.4. Los datos.....	pág.33

### **CAPÍTULO 4: LOS VERBOS IRREGULARES EN ESPAÑOL**

4.1. Cuándo un verbo se considera irregular?.....	pág. 36
4.1.1. ¿Cómo se manifiestan las irregularidades?.....	pág.37
4.2. Cambios vocálicos en la raíz.....	pág. 38
4.2.1. Cambios por diptongación.....	pág. 39
4.2.2. Cambios por cierre de vocal.....	pág. 39
4.3. Cambios consonánticos.....	pág. 40
4.4. Irregularidades especiales.....	pág. 40
4.4.1. Pretéritos fuertes.....	pág. 41
4.5. Consideraciones finales.....	pág.42

### **CAPÍTULO 5: LOS VERBOS IRREGULARES EN LA INTERLENGUA DE APRENDIENTES DE ESPAÑOL**

5.1. Clasificación de los verbos empleados en la interlengua.....	pág. 43
5.2. Verbos irregulares empleados según la L1.....	pág. 50

### **CAPÍTULO 6: FORMAS VERBALES DESVIADAS**

6.1. Clasificación de las formas verbales desviadas.....	pág. 53
6.2. Formas verbales desviadas por corpus.....	pág. 54
6.2.1. Corpus francés.....	pág. 54
6.2.2. Corpus inglés.....	pág. 56
6.2.3. Corpus portugués.....	pág. 57
6.2.4. Corpus alemán.....	pág. 59

6.2.5. Corpus holandés.....	pág. 61
6.3. Consideraciones finales.....	pág. 63

<b>2DA PARTE: LINGÜÍSTICA COMPUTACIONAL.....</b>	<b>Pág. 65</b>
--	----------------

## **CAPÍTULO 7: TÉCNICA ESTADÍSTICA NO PARAMÉTRICA EN LA COMPARACIÓN DE CORPUS DE DISTINTAS LENGUAS DE ORIGEN**

7.1. Herramienta informática: Planilla Excel.....	pág. 67
7.2. Técnica estadística no paramétrica en la comparación de corpus de distintas lenguas....	pág. 78
7.2.1. Resultados.....	pág. 81
7.3. Frecuencia de verbos.....	pág. 84
7.4. Resultados.....	pág. 95
7.5. Conclusiones.....	pág. 101

## **CAPÍTULO 8: HERRAMIENTAS INFORMÁTICAS.....**

8.1. Smorph.....	..pág.105
8.1.1. Reconocimiento de palabras a partir de terminaciones distinguidas.....	..pág.108
8.1.2. Análisis automático de Smorph.....	...pág.110
8.2. Sistema Nooj.....	...pág.113
8.2.1. Análisis morfológico de Nooj.....	...pág.117
8.2.2. Localización de términos en Nooj.....	..pág. 122

## **CAPÍTULO 9: TRATAMIENTO AUTOMÁTICO PARA LOS VERBOS IRREGULARES DE ESPAÑOL ESTÁNDAR.....**

9.1. Modelos verbales irregulares de Smorph.....	pág. 124
9.1.1. Ejemplos de modelos verbales irregulares.....	pág. 127
9.1.2. Verbos con cambios gráficos.....	pág. 128
9.1.3. Verbos hiperirregulares.....	pág. 130
9.2. Modelos verbales en el sistema Nooj.....	.pág.134

## **CAPÍTULO 10: ANÁLISIS AUTOMÁTICO DE LA INTERLENGUA.....**

pág. 137

10.1. Implementación en Smorph.....	.pág.137
10.1.1. Modificaciones en el archivo Entradas.....	pág. 142
10.1.2. Modificaciones en el archivo Rasgos .....	pág. 143
10.1.3. Modificaciones en el archivo Modelos.....	pág. 143
10.1.4. Resultado del análisis de Smorph.....	pág. 147
10.2. Implementación en Nooj.....	pág. 149
10.2.1. Generación de gramáticas productivas.....	. pág.149
10.3. Conclusiones.....	pág.160
<b>CONCLUSIONES.....</b>	Pág. 165
Etapas de la investigación.....	pág.165
Evaluación de las hipótesis lingüísticas.....	pág.167
Evaluación de hipótesis en el aspecto informático.....	pág.168
Aportes de mi tesis.....	pág.170
Proyecciones.....	pág.170
<b>BIBLIOGRAFÍA.....</b>	pág.171
<b>ANEXOS.....</b>	PÁG.176
Anexos Nº1 Corpus producciones de aprendientes de español.....	pág.176
Anexos Nº2 Tablas dinámicas.....	pág.200
<b>Agradecimientos.....</b>	pág.233

La presente investigación se enmarca por un lado en la lingüística informática y por otro, en la adquisición de L2. La iniciativa busca cubrir un lugar de vacancia, dada la escasez de estudios sobre adquisición de L2 dentro de la lingüística informática.

### Lingüística computacional

En los últimos tres décadas surge la Lingüística Computacional como ciencia del lenguaje que contribuye al conocimiento de los procesos cognitivos de comprensión y producción del lenguaje, combinando teorías del lenguaje natural con técnicas estadísticas. Es un área interdisciplinaria que toma saberes de la Lingüística y la Informática y su tarea consiste en crear sistemas informáticos capaces de procesar el lenguaje humano.

Siguiendo a Lavid (2005) el objetivo que se propone es el de diseñar programas informáticos que puedan emular en vez de simular la capacidad lingüística humana. Emular no significa comprender cómo funciona el cerebro humano si no intentar construir sistemas que comprendan y produzcan el lenguaje de manera similar a un humano.

En la Universidad Nacional de Rosario, desde el año 2004 se dictan seminarios de Lingüística computacional y en 2006 se crea el grupo INFOSUR<sup>1</sup> que se encuentra en ejecución en el Centro de Estudios de Adquisición del Lenguaje bajo la dirección de Zulema Solana. Es importante mencionar que a raíz de los estudios llevados a cabo en el marco del proyecto se han desarrollado y aprobado cinco tesis de Doctorado sobre análisis automático, entre ellas la investigación llevada a cabo por Celina Beltrán (2008) que realiza la modelización lingüística e informática estadística en el análisis automático de textos, la de Rodolfo Bonino (2010) sobre construcciones causativas, la de Andrea Rodrigo (2012) sobre el sintagma adverbial núcleo, la de Walter Koza (2012) sobre los signos de puntuación y la de Bárbara Méndez (2013) sobre el sintagma adjetivo núcleo.

---

<sup>1</sup> INFOSUR (investigación y desarrollo)

Este trabajo se basa fundamentalmente en los proyectos de investigación del grupo INFOSUR ya que gracias a la implantación de las categorías y los modelos flexivos del español, que se ha efectuado en Smorph<sup>2</sup>, actualmente el grupo de trabajo se encuentra trabajando en el traslado de esa información lingüística de acuerdo a los requerimientos del sistema Nooj<sup>3</sup>.

## Objetivos

Me propongo adaptar la formalización lingüística de los programas informáticos (que fueron diseñados para el reconocimiento de lenguas naturales) con la finalidad de poder efectuar el análisis automático de las formas verbales irregulares halladas en textos de aprendientes de español como segunda lengua.

### Objetivo generales:

Realizar la implantación en máquina de la morfología verbal irregular a fin de analizar automáticamente textos de aprendientes de español.

-Contribuir al estudio del análisis automático de la interlengua y del proceso de adquisición del español como segunda lengua.

### Específicos:

.Realizar una adaptación de la información lingüística declarada en los archivos que componen las herramientas informáticas Smorph y Nooj para lograr detectar formas verbales irregulares propias de la interlengua.

-Comparar los resultados obtenidos analizando los alcances de dos herramientas informáticas diferentes.

- Evaluar las hipótesis lingüísticas propuestas sobre la adquisición de la flexión verbal irregular.

---

<sup>2</sup> desarrollado en el GRIL por Salah Aït Mokhtar (1998)

<sup>3</sup> Es un sistema para el tratamiento de lenguas naturales creado por Max Silverztein (2002) en la Universidad de Franche-Comté que se encuentra disponible on line para ser descargado y utilizado.

## Fundamentación

La elección del tema surge como consecuencia de inquietudes personales nacidas de la situación de enseñanza del español como segunda lengua a extranjeros. Sin embargo, el objetivo general de la investigación no es didáctico sino que focaliza la adquisición y pretende analizar la interlengua en su totalidad, teniendo en cuenta tanto las formas desviadas respecto de la lengua estándar como las coincidentes con el sistema, siendo su finalidad poder ejecutar, en la medida que esto sea posible, el análisis automático de la interlengua.

Página | 8

## Estado de la cuestión

En lo que refiere a la adquisición de L2, en la década del 40 del siglo pasado, en el ámbito de la lingüística aplicada, comienzan las investigaciones que tienen como objetivo el estudio de la lengua del aprendiente. Se trata de predecir, mediante una gramática contrastiva, las dificultades con las que se encontrará el estudiante para poder evitar el error. A partir de la teoría de la Gramática Universal propuesta por Chomsky en 1959, surge la corriente del Análisis de Errores con la distinción que realiza Corder (1967) entre error y falta.

En la década del 90 se crean varias clasificaciones de errores. Para la enseñanza del español como segunda lengua tenemos por ejemplo, la de Sonsoles Fernández (1995). Las mismas centran su atención en las desviaciones respecto de la norma de la lengua meta. Con los estudios de la interlengua en cambio, el investigador considera toda la producción del aprendiz, es decir, tanto la correcta como la incorrecta. El término es utilizado por Selinker en 1972, proponiendo un enfoque psicolingüístico y definiendo a esta lengua en un lugar intermedio entre la lengua materna y la meta, teniendo elementos comunes con ambas.

Mi trabajo se basa en dicha concepción de la interlengua y contempla en su conjunto tanto a las formas coincidentes como a las formas idiosincrásicas que presenta este sistema.

Los trabajos que han representado un aporte en lo que respecta a este ámbito son los de Alexopolou (2005) que clasifica los errores según su causa sea intralingüística o interlingüística, Rigamonti (2006) acerca de los problemas en la adquisición, entre otros.



En el campo de la lingüística computacional, se reconocen tres líneas de investigación: estadística, lingüística y una combinación de ambas. Dentro de este primer enfoque, pueden mencionarse los trabajos de Dister (1997) y Reynar y Ratnaparkhi (1997).

Con respecto a la última línea cabe mencionar la propuesta de Celina Beltrán (2009), que plantea una comparación de los métodos utilizados en la modelización lingüística y métodos fundados en enfoques estadísticos.

En este estudio introduzco algunas técnicas estadísticas básicas porque considero que el trabajo de análisis lingüístico automático es enriquecido notablemente si se combina con la estadística, sobre todo al utilizar un corpus que reúne variadas lenguas de origen.

### Hipótesis

Las hipótesis de trabajo que desarrollaré son las siguientes:

1) En el aspecto informático:

Se pueden analizar automáticamente las producciones de aprendientes de español, logrando el reconocimiento de formas propias de la interlengua, si se modifica la información lingüística declarada en los archivos de las herramientas informáticas de las que dispongo, de acuerdo a las características peculiares que presenta la interlengua.

2) En lo que hace a la interlengua:

Se puede predecir, a partir de las observaciones realizadas hasta el momento, que:

b) Respecto de la morfología flexional, se encuentran diferentes tipos de desviaciones en la conjugación de verbos irregulares.

Puedo arriesgar, a partir de las observaciones y el análisis de los datos realizado hasta el momento, que existe una clara tendencia a que las desviaciones en el nivel morfológico flexivo respondan a fenómenos y características particulares del español, es decir, que la causa sería intralingüística y por lo tanto, las mismas desviaciones generadas en formas verbales irregulares se darían en las producciones de aprendientes de nivel inicial que poseen como lengua de origen el alemán, francés, inglés, holandés y portugués.

Para comprobarlo se emplean técnicas estadísticas que determinen qué cantidad de formas verbales irregulares son coincidentes con la lengua estándar y qué cantidad presenta desviaciones respecto a ésta.

La hipótesis central de la investigación radica entonces en que el fenómeno recurrente de hallar formas verbales irregulares que no son coincidentes con las de la lengua meta se presenta en todos los estratos en la misma proporción, es decir, en los corpus compuestos por las diferentes lenguas de origen de los aprendientes de español.

### Etapas y Metodología de trabajo:

Las etapas que comprende la investigación son las siguientes:

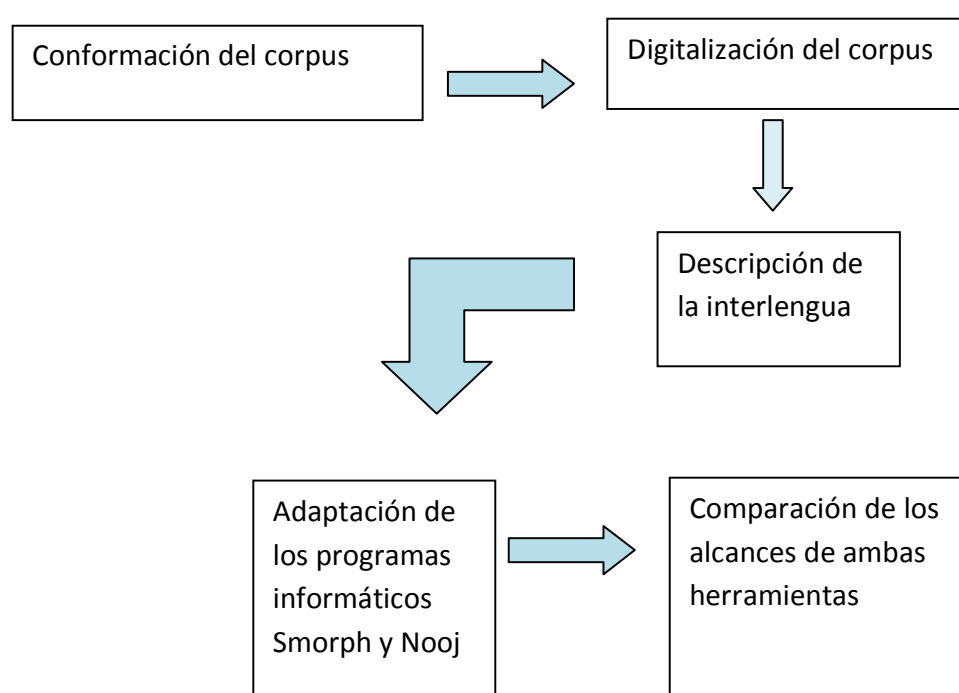


Gráfico N1 Etapas de la investigación

En primer lugar, reuní un corpus considerable de producciones de aprendientes de español como segunda lengua, que se encuentran en un nivel inicial de adquisición y

poseen variadas lenguas de origen, a saber: alemán, inglés, francés, portugués, holandés. En total, cuento con un total de 140 producciones y 13000 palabras.

En segundo lugar, fue necesaria la digitalización del corpus, es decir la transcripción de los textos escritos, requisito indispensable para comenzar el análisis automático realizado por los programas informáticos.

A partir de esto, realizo una descripción de la interlengua tomada como corpus y observo un fenómeno recurrente: desviaciones respecto de la lengua estándar en formas verbales irregulares.

La descripción lingüística de estas formas idiosincrásicas encontradas implicó un estudio de los verbos irregulares, siguiendo la investigación de Alcoba Santiago (1999) explico cuándo un verbo se considera irregular, cuáles son las irregularidades que presenta el sistema verbal español, qué verbos que presentan más de un tipo de irregularidad al flexionarse, entre otras cuestiones.

La metodología de trabajo implica un estudio cuantitativo en cuanto al empleo de pruebas estadísticas para medir el porcentaje de desviaciones y de formas coincidentes pero además un estudio cualitativo ya que requiere la descripción de los datos para su posterior clasificación y análisis.

Teniendo en cuenta la evaluación de la cobertura y precisión, la última etapa de la investigación consiste en comparar los alcances de dos programas informáticos en cuanto al reconocimiento de la morfología verbal irregular de la interlengua.

Por consiguiente, describo el funcionamiento de los dos sistemas y las ventajas con las que cuentan uno y otro en cuanto al reconocimiento automático de las formas verbales irregulares desviadas reveladas en el corpus.

## Herramientas

En cuanto a las herramientas, utilizo Excel para confeccionar tablas que funcionen como base de datos y a partir de las cuales pueda obtener el total de ocurrencias y el total de desviaciones tanto en cada lengua en particular como en todos los corpus en general.

Además empleo la técnica estadística no paramétrica de Kruskal Wallis para determinar qué cantidad de formas verbales irregulares son coincidentes con la lengua estándar y qué cantidad presenta desviaciones respecto a ésta.

Con el fin de establecer cuáles de los verbos irregulares han sido más utilizados y cuáles de ellos presentan una mayor cantidad de desviaciones en la interlengua confecciono una tabla dinámica en Excel.

Para cumplir el objetivo de lograr el análisis automático de la interlengua en su totalidad, incluidas las formas verbales que presentaban alteraciones respecto al español estándar, es necesario realizar adaptaciones en los archivos correspondientes a las herramientas informáticas con las que trabajo. Particularmente utilizo Smorph que es un analizador y generador morfosintáctico desarrollado en el GRIL por Salah Aït Mokhtar (1998) y Nooj que es un sistema para el tratamiento de lenguas naturales creado por Max Silverztein (2002) en la Universidad de Franche-Comté que se encuentra disponible on line. Es importante mencionar que el grupo de trabajo INFOSUR cuenta con el asesoramiento y la supervisión de Max Silverztein, quién dictó seminarios en la Universidad Nacional de Rosario en el año 2015.

#### **APORTES DE LA TESIS**

La originalidad de la propuesta radica en poder emplear programas informáticos creados para el procesamiento del lenguaje de lenguas naturales para analizar textos de la interlengua de español como L2.

Considero que el aporte de mi tesis consiste en abrir el campo de estudio en lo que respecta al tratamiento informático de la interlengua, pudiendo contribuir a los estudios de adquisición del español y de la enseñanza de segundas lenguas mediante la utilización de herramientas informáticas que sirvan como recursos didácticos.

#### **ORGANIZACIÓN**

La tesis está dividida en tres partes:

En la primera parte se explicitan ciertas cuestiones teóricas concernientes a la adquisición de segundas lenguas. Se coloca en primer lugar el capítulo sobre la teoría de la interlengua, en éste se desarrolla el concepto de interlengua atendiendo a sus características y se explica la naturaleza de los errores según respondan estos a una causa intralingüística o interlingüística.

En el capítulo dos se realiza un breve recorrido a través de los modelos científicos surgidos a partir de la segunda mitad del siglo XX en el marco de la adquisición de segundas lenguas: el análisis contrastivo y el análisis de errores. Al final de éste se incluye un glosario de términos propios del campo de adquisición de segundas lenguas.

El tercer capítulo explica la conformación y organización del corpus, además escribe el origen de los sujetos que forman parte de la prueba y muestra los tipos de consignas y ejercicios que han servido para la recolección de las producciones escritas.

El cuarto capítulo trata sobre la flexión verbal irregular del español distinguiendo los cambios vocálicos de los cambios consonánticos y designando a un grupo de verbos como de irregularidades especiales.

La segunda parte, correspondiente al análisis de los datos, contiene el capítulo cinco que describe los verbos irregulares hallados en el corpus de interlengua y realiza una clasificación de éstos según la variación que presenten.

En el capítulo seis se agrupan las desviaciones observadas en cuatro casos y se muestran ejemplos de estas formas idiosincrásicas en cada corpus según la L1.

La tercera parte corresponde al tratamiento computacional. Comienza con el capítulo siete en donde se realiza un análisis estadístico de los datos, calculando el porcentaje de desviaciones y de formas coincidentes con la lengua estándar en cada uno de los corpus. Continúa con el capítulo ocho en el que se describe el funcionamiento de las herramientas informáticas haciendo hincapié en los modelos verbales que se han creado para analizar automáticamente los verbos irregulares.

El capítulo noveno refiere al tratamiento automático de verbos irregulares en los dos programas utilizados: Smorph y Nooj.

En el capítulo décimo se muestran las modificaciones efectuadas en los programas para el reconocimiento de los cuatro casos de formas verbales irregulares desviadas de la interlengua: los cambios en los archivos de Smorph y la creación de gramáticas productivas en Nooj. Se miden los resultados obtenidos comparando los alcances de ambas herramientas según la cobertura y precisión.

Para finalizar, se enuncian las conclusiones que incluyen los aportes de la tesis y las proyecciones

## CAPÍTULO 1: LA TEORÍA DE LA INTERLENGUA

En este capítulo se desarrolla el concepto de interlengua atendiendo a sus características. Se analiza la incidencia de la lengua materna en la adquisición de la morfología y se explica la naturaleza de los errores según respondan estos a una causa intralingüística o interlingüística. Tal distinción permitirá establecer una primera hipótesis respecto del origen de las desviaciones halladas en la flexión de verbos irregulares en el corpus analizado.

El campo de investigación de la adquisición de segundas lenguas tuvo un auge considerable en las décadas de 1950 y 1960 con la hipótesis del análisis contrastivo, que se encargó de describir y comparar las lenguas involucradas en el proceso (la lengua nativa del aprendiente y la lengua meta) tratando de hallar los puntos de contacto y las zonas divergentes entre uno y otro sistema (ya sea en cuanto a fonología, morfología, y sintaxis) con el objetivo de detectar y predecir los errores que cometerían los estudiantes.

Al advertir que la fuente de la causa de los errores no sólo respondía a la interferencia<sup>4</sup> con la lengua materna (entendida ésta como el efecto de la lengua nativa sobre la lengua meta) surge un nuevo paradigma: el del análisis de errores. El aporte de esta línea investigativa reside en despojar de negatividad al error y tomarlo como un indicador del proceso de aprendizaje llevado a cabo por el aprendiente.<sup>5</sup>

Más tarde, en la década de 1970 se llega al desarrollo de la Teoría de la Interlengua que significa un giro metodológico al proponer el estudio y análisis tanto de las producciones desviadas como de las correctas. Esta teoría toma en consideración la producción total de los estudiantes, demostrando que tanto unas como otras son importantes en el proceso de aprendizaje.

<sup>4</sup> Ver índice de conceptos al final del capítulo Nº 2.

<sup>5</sup> En el capítulo siguiente se ampliará la descripción de los mencionados estudios de adquisición de segundas lenguas.

## 1.1. Concepto de interlengua

El término interlengua es utilizado por Selinker en 1972, proponiendo un enfoque psicolingüístico y colocando a esta lengua en un lugar intermedio entre la lengua materna y la meta, teniendo elementos comunes con ambas:

*Cuando formulemos los principios relevantes para una teoría del aprendizaje de lenguas segundas estaremos (...) obligados a presentar como hipótesis la existencia de un sistema lingüístico independiente: sobre la base de los datos observables que resultan de los intentos del alumno en la producción de una norma de la lengua objeto. Llamaremos “**Interlengua**”(IL) a este sistema lingüístico.<sup>6</sup>*

Veamos ahora cómo influye la teoría del análisis de errores en esta definición:

Para dicha teoría los errores son los que caracterizan a la lengua del aprendiente, que se manifiesta como un sistema lingüístico distinto tanto al de la L1 como al de la L2.

Corder (1967) denomina a este sistema como un *dialecto idiosincrásico* o *dialecto transitorio* y le atribuye las siguientes características: es un sistema lingüístico con una gramática propia que se encuentra en continuo cambio, sus oraciones son indiosincrásicas, no erróneas y es un sistema peculiar de un estudiante individual o de un grupo de estudiantes que posee la misma formación académica.

Estas etapas o niveles de competencia se van modificando a medida que el aprendiz va adquiriendo vocabulario y nuevas estructuras de la lengua meta.

Es interesante señalar de esta perspectiva el suponer que aunque esta lengua idiosincrásica difiera en cada alumno en particular, presenta zonas de intersección en aprendientes con un nivel académico similar e idéntica lengua nativa.

Nemser (1969) se refiere al mismo concepto pero denominándolo *sistema aproximativo* para caracterizar al sistema lingüístico desviado que emplea el aprendiente al tratar de utilizar correctamente la lengua que está aprendiendo, atribuyéndole también una gramática y vocabulario propios.

---

<sup>6</sup> El subrayado es mío. Traducción de Liceras (2005:90)

De estas definiciones podemos extraer propiedades comunes: en primer lugar el aprendiz posee un sistema, considerado *dialecto*, *competencia* o *lengua*, otra característica es que se encuentra en continuo cambio, que no es estable, de ahí los calificativos: *transitorio*, *aproximado* e *inter*, además de ser propio de los que aprenden una nueva lengua y contar con reglas específicas: *idiosincrásico*.

## 1.2. Características de la interlengua

Con los estudios de la interlengua entonces, el investigador considera toda la producción del aprendiz, es decir, tanto la correcta como la incorrecta.

La relevancia de los estudios de la interlengua, coincidiendo con Alexopoulou (2010: 2)

“... surge de la necesidad de descubrir los principios generales que determinan el aprendizaje de la lengua extranjera, según los cuales el cerebro humano de los aprendientes procesa los datos del lenguaje a los cuales están expuestos.”

Es importante aclarar que la interlengua contiene reglas propias y no debe tomarse como una mezcla entre los dos sistemas lingüísticos. La literatura especializada le otorga dos características que sin embargo parecen contradictorias: sistematicidad y variabilidad.

La primera refiere al conjunto de reglas coherente con los que los aprendientes realizan hipótesis, sin embargo éstas son variables, se modifican a medida que el estudiante avanza en la adquisición y pasa a otra etapa. A partir del nuevo estadio se reestructuran esas reglas e hipótesis.

La **sistematicidad** entonces está dada por la coherencia interna que presenta esa interlengua en un determinado momento de su desarrollo, se vislumbra en la aplicación de reglas lingüísticas que responden a estrategias y procesos que activan los aprendientes. En esta gramática singular, encontraremos oraciones o construcciones “desviadas” que sin embargo, serán adecuadas desde el punto de vista del aprendiente.

Es importante destacar que ese sistema lingüístico tiene características y peculiaridades propias que coinciden con las halladas en el sistema de todo estudiante que está adquiriendo una segunda lengua, si se toma en cuenta un mismo nivel de aprendizaje.

Por lo tanto, en esta investigación se comprobará la existencia de reglas lingüísticas a la hora de flexionar verbos irregulares en la interlengua de aprendientes de español con distintas lenguas maternas (alemán francés, holandés, inglés y portugués).



Con respecto al atributo de **variabilidad**, éste responde al carácter transitivo de la interlengua ya que es un sistema que se encuentra en constante reestructuración, en un proceso de cambios, en el camino hacia la adquisición de la lengua meta. Podemos decir que se halla en continua transformación y que da lugar a un pasaje por distintos estadios o etapas, las cuales poseen sus propias reglas, hasta alcanzar el estado final.

Ejemplos de variabilidad son los casos en que el aprendiente emplea dos o más variantes para expresar un fenómeno lingüístico que posee una sola realización.

Como ejemplo, encontramos que en un mismo texto se utiliza en cinco oportunidades la palabra *ciudad*. El estudiante (cuya lengua materna es el portugués) emplea cuatro veces la forma coincidente con español: *ciudad* y en una ocasión usa *cidade*<sup>7</sup>:

*Quando llegei em Rosario, 16 de janeiro de 2012, me senti surpreso com as novidades. O tempo estava muito agradável. Muitas coisas eram diferentes da minha **cidade**. Havia muitos cães pelas ruas da **cidade**. Também havia muitos casais de idosos por toda a **cidade**. Senti muitas diferenças e as pessoas falavam muito pouco. Eles não sorriam, pareciam não gostar muito dos brasileiros. Agora, julho de 2012, por segunda vez, me senti mais sereno. Vi muito poucos cães pelas ruas. A **cidade** parecia muito mais bonita mas o tempo muito frio. Os professores pareciam mais satisfeitos. Também o edifício da universidade está mais confortável. A **cidade** parece que está com muito mais pessoas. Cada vez que estou me sentindo como um rosário. Estou me sentindo em minha própria casa. Quando terminar meu doutorado, quero voltar outras vezes em Rosario.*

### 1.3. ¿Cómo incide la lengua materna en la adquisición?

Numerosas investigaciones plantean la existencia de una secuencia natural de adquisición que es independiente de la lengua materna del aprendiz, esto implicaría un orden y un ritmo universal dentro de cada etapa, como la de Dulay y Burt (1973) que investigan la

<sup>7</sup> Ver anexo corpus portugués

adquisición de marcadores ingleses en niños hispanohablantes y chinos, concluyendo que existe un mismo orden de adquisición para ambos grupos de aprendientes.

Adjémian (1982) considera que la idealización realizada por Chomsky para la adquisición de L1 es válida para analizar la interlengua, teniendo en cuenta el hecho de que los adultos recurren a las capacidades metalingüísticas y a la lengua materna para construir la gramática de interlengua. (citado por Baralo, 2004).

Selinker propone la existencia de una estructura psicológica latente, tomando el concepto de *estructura latente del lenguaje* de Lenneberg (1967), la cual es activada por el aprendizaje de L2 al momento de expresarse en una determinada situación comunicativa. En esta estructura psicológica se dan las *identificaciones interlingüísticas*, que constituyen el lazo entre los sistemas: la LM, la L2 y IL.

Para este autor, una cuestión importante para los estudios de interlengua, es el concepto de fosilización, con éste se refiere a la permanencia en la IL de ciertas reglas gramaticales, vocabulario o aspectos de pronunciación, más allá de la cantidad de instrucción recibida o de “input”. Observa que los fenómenos fosilizables permanecen en la IL aún a través del paso del tiempo y por eso sostiene que esas estructuras psicolingüísticas se hallan en el cerebro y se encuentran almacenadas por un mecanismo de fosilización.

Los principales procesos de fosilización, que puntualiza Alexopoulou (2005: 92, 2010) son:

1. La transferencia lingüística: fosilización como resultado de la influencia de la LM.
2. La transferencia de instrucción: la fosilización tiene como causa la aplicación de una técnica didáctica inadecuada.
3. Las estrategias de aprendizaje de la segunda lengua: reflejan los procesos internos que el alumno pone en marcha para aprender la información lingüística. Por ejemplo: la memorización de las nuevas estructuras.
4. Las estrategias de comunicación en la segunda lengua: cuando la fosilización se produce en los intentos del alumno por comunicarse en la lengua meta al no poseer una competencia suficiente.
5. La hipergeneralización del material lingüístico de la L2: la fosilización es producto de una falsa regularización de reglas de la lengua meta.

Un ejemplo de este último proceso es la flexión de verbos irregulares en la adquisición del español como segunda lengua. Particularmente, en la regularización de las formas verbales irregulares (*quieren* en vez de *quieren*, *encontra* en vez de *encuentra*), que se agrupan como desviación 1, grupo de desviaciones que presenta la mayor cantidad de casos en todos los corpus.

#### 1.4. Clasificación de errores según su causa

Es importante distinguir dos tipos de interferencia, si se tiene en cuenta el efecto de la lengua meta en la nativa, como lo consideró el posterior análisis de errores, se habla de *Interferencia intralingüística* (interna), es decir, cuando la lengua que se aprende es la responsable de los errores formulados en la producción lingüística.

Por el contrario, cuando la lengua madre es la causa de los errores en la producción lingüística se habla de *Interferencia Interlingüística* (externa).

##### 1.4.2. Errores intralingüísticos

Una de las causas entonces de los llamados errores producidos por los aprendientes es la hipergeneralización y aplicación incompleta de las reglas de la lengua meta. Refiere a las hipótesis que realiza el estudiante que está aprendiendo otra lengua, basándose en el conocimiento que posee de la estructura o del funcionamiento de la lengua meta, el resultado de este fenómeno son los errores intralingüísticos.

Tomando la clasificación que realiza Vásquez (1991) según la causa del error, es decir, utilizando un criterio etiológico, tenemos:

Errores intralingüísticos, que dependen de la L2, por ejemplo regularizar un participio irregular : hemos hacido (del corpus inglés), regularizar un verbo irregular: *empezo* .

Errores interlingüísticos: dependen de la L1 o de otra lengua aprendida: por ejemplo la terminación verbal –ei en formas como *desembarquei*, propia de la lengua portuguesa observada en producciones del corpus portugués. El empleo de doble consonante (ss), por ejemplo en el corpus alemán.

Errores de simplificación: si podía lo hacía

La hipótesis que sostengo en la presente investigación es que ésta es la causa de las desviaciones encontradas en los diferentes corpus en lo que atañe a la flexión de verbos irregulares.

Por lo tanto, regularizar formas que son irregulares, hacer irregular una forma que no varía, y colocar desinencias existentes en el español estándar pero que no corresponden a

esas bases verbales responden a la aplicación de reglas de forma generalizada para conjugar los verbos que evidencia que todos los estudiantes (de distintas lenguas de origen) atraviesan la misma etapa en el proceso de aprendizaje del español.

La metodología de trabajo atenderá tanto a un estudio cuantitativo como cualitativo, ya que analiza la recurrencia numérica de desviaciones pero a la vez describe, explica y clasifica a éstas para llegar a un conocimiento profundo de la problemática.

Se comprobará entonces si esta dificultad para flexionar verbos irregulares que presentan distintas variaciones a lo largo de su conjugación y que responden a una particularidad del sistema verbal español se halla en la misma proporción en todos los corpus trabajados.

En este capítulo se realiza un breve recorrido a través de los modelos científicos surgidos a partir de la segunda mitad del siglo XX en el marco de la adquisición de segundas lenguas<sup>8</sup>. El objetivo es dar cuenta de los aportes realizados por el análisis contrastivo y el análisis de errores en la construcción de la teoría sobre estudios de interlengua que predomina en la actualidad<sup>9</sup>. Para finalizar y a modo de compendio se establece una comparación entre las tres líneas teóricas y se incluye un glosario de términos referidos al área lingüística que atañe a la adquisición de L2.

### 2.1. El análisis contrastivo

La Hipótesis del Análisis Contrastivo, que se desarrolló durante las décadas de los cincuenta y los sesenta está fundamentada en las investigaciones de C. Fries (1945) y R. Lado (1957). Ésta se encarga de describir las diferencias y las similitudes de dos sistemas lingüísticos, es decir de la lengua materna y de la segunda lengua, en todos los niveles de análisis del lenguaje.

El objetivo es poder determinar qué áreas o estructuras lingüísticas traerán problemas a la hora del aprendizaje para poder predecir, a partir de las diferencias entre ambas lenguas, qué tipo de errores cometerán los aprendientes.

Se trata de pronosticar, mediante la creación de una gramática contrastiva, las dificultades con las que se encontrará el estudiante con el fin de evitar el error. Su preocupación es didáctica.

---

<sup>8</sup> De ahora en más utilizo L2 para referirme a *segunda lengua*.

<sup>9</sup> Cabe aclarar que a lo largo de la investigación se emplearán los términos aprendizaje y adquisición indistintamente para referir al proceso de apropiación por parte del aprendiz de la lengua meta.

Como afirma Lado, R. (1957) en el Prefacio de su libro: “El plan de este libro descansa en la afirmación de que podemos predecir y describir las estructuras que causarán dificultad en el aprendizaje, y aquellas que no entrañarán dificultad, mediante la comparación sistemática de la lengua y cultura que va a ser aprendida con la cultura y la lengua nativa del estudiante” (R. Lado, 1957, citado de I. Gargallo (2006).

Desde el análisis lingüístico sigue los siguientes pasos (Sonsoles Fernández: 1995):

- 1) Descripción formal de los idiomas en cuestión,
- 2) selección de las áreas que van a ser comparadas,
- 3) comparación de las diferencias y semejanzas
- 4) predicción de los posibles errores.

Este tipo de análisis, como explica Rigamonti (2006), está basado en la teoría psicológica del aprendizaje del conductismo: ante estímulos que requieran respuestas positivas o negativas, se consolidan ciertos hábitos y se rechazan otros. Trasladado este modelo a la enseñanza de una lengua, al estudiante se le presentan ejercicios de repetición o completamiento que no requieren de una total comprensión del contexto o que no implican un razonamiento por parte del éste. El profesor interviene para prevenir posibles errores reforzando hábitos lingüísticos a través de la repetición y la mecanización de las estructuras lingüísticas aprendidas.

El aprendizaje se concibe entonces como un proceso de imitación de un modelo correcto (establecido a través del profesor o del libro de texto) y el objetivo es adquirir los hábitos adecuados en la segunda lengua.

Esta corriente sostiene que los errores de los que aprenden una L2 son el resultado de la interferencia con su lengua materna, generada por las diferencias entre las estructuras de los dos sistemas lingüísticos. A esta interferencia también se la llama transferencia negativa.

En cambio, cuando las estructuras de las dos lenguas en cuestión presentaran semejanzas, el efecto sobre el aprendizaje sería positivo y la transferencia, por lo tanto, positiva.

Por lo tanto, cuanto más disímiles fueran las lenguas en cuestión, mayores dificultades habría en el aprendizaje por existir mayor probabilidad de interferencia.

R. Wardhaugh (1970)<sup>10</sup> hace una distinción entre hipótesis fuerte e hipótesis débil. En la versión más fuerte se predice mediante la comparación mientras que en la débil se diagnostica y se explican las desviaciones que han sido halladas.

Entre las principales críticas negativas que recibe este modelo encontramos que no logra evitar los errores y que la mayoría de éstos no se deben a la interferencia con la lengua materna sino que están relacionados con factores internos de la L2.

Ocurre que más tarde se advierte que los errores producidos por los aprendientes no tienen como única causa la interferencia con su primera lengua, ya que estudios empíricos realizados dieron como resultado que no más del 33 % de los errores se debía a una interferencia interlingüística.

Esto permite el surgimiento de una nueva línea científica: la del Análisis de Errores que trabaja analizando los errores manifestados en la producción oral y escrita de los sujetos que aprenden una segunda lengua.

A pesar de las críticas recibidas es importante destacar del AC el hecho de haber constituido el punto de arranque de las posteriores investigaciones en el campo de la adquisición y el origen de los estudios actuales sobre Interlengua.

## 2.2. Análisis de Errores

A partir de la teoría de la Gramática Universal propuesta por Chomsky en 1959, surge la corriente del Análisis de Errores.

Chomsky (1989) plantea que el individuo nace con una capacidad innata para adquirir el lenguaje. Le otorga una participación activa al aprendiz ya que éste es el encargado de realizar hipótesis sobre el funcionamiento de la lengua en la que entra en contacto.

El aporte que realiza el modelo teórico del Análisis de Errores es el de considerar a éstos como indicios de que el proceso de aprendizaje se está llevando a cabo.

A diferencia de lo que se pensaba en el Análisis contrastivo, se buscan las causas de los errores en la lengua meta y no en la lengua materna ya que se observan los mismos errores en aprendientes de distintas lenguas.

<sup>10</sup> Wardhaugh distingue las dos versiones en su artículo "The contrastive analysis hypothesis" (traducción en Liceras 1991)

Corder (1967) realiza una distinción entre error –ligado al concepto de competencia de Chomsky – y falta –ligado al de actuación -. La competencia abarca mucho más que la actuación, es la capacidad que tiene el individuo de expresarse y por lo tanto es infinita. La actuación refiere a la expresión concreta del lenguaje. El error entonces es algo que comete siempre el aprendiz, una desviación sistemática de la norma de la lengua meta, mientras que la falta se debe a razones que no están ligadas al conocimiento de la lengua, son desviaciones asistemáticas, que pueden deberse a una distracción del estudiante, por lo tanto es factible que lo advierta y se autocorrija.

Corder considera que los errores son un fenómeno inevitable en el proceso de aprendizaje de lenguas, con respecto a los estudios contrastivos cree que debería emplearse para confeccionar un inventario de errores más frecuentes.

A diferencia del modelo anterior ya no se someten a comparación las dos lenguas sino que se consideran las producciones reales de los aprendientes y luego se sigue el procedimiento de identificar los errores en su contexto, clasificarlos y describirlos, explicarlos mediante las estrategias psicolingüísticas y las causas de cada error y por último si el objetivo es didáctico, evaluar la gravedad de los errores y buscar cómo remediarlos. (Corder 1967, ed.1981, 14-26)

En la década del 90 se crean varias clasificaciones de errores. Éstas centran su atención en las desviaciones respecto de la norma de la lengua meta. Cada una de ellas se presenta como una taxonomía que tiene en cuenta diferentes criterios para la clasificación de los errores.

Para la enseñanza del español como segunda lengua tenemos por ejemplo, la de Sonsoles Fernández (1997). Esta investigación abarca todo el sistema de la lengua: léxico, morfosintaxis, discurso y los errores fonológicos que revela la grafía de producciones de aprendientes de español con cuatro lenguas maternas diferentes y en tres niveles de competencia, proponiéndose extraer un corpus de errores (universales o característicos) de cada grupo de lengua materna.

Debido a que los errores se encasillan de manera no natural en las distintas etiquetas que son creadas para poder clasificarlos, este análisis de errores se convierte en una descripción sin frutos, por lo menos en cuanto a aplicaciones didácticas respecta. Su máxima contribución es la de tener como foco de atención al error pero considerándolo positivo para el aprendizaje y no como algo censurable.



### 2.3. Consideraciones finales

El siguiente cuadro compara los tres modelos de análisis en cuanto al aprendizaje de una lengua extranjera:

	ANÁLISIS CONTRASTIVO	ANÁLISIS DE ERRORES	INTERLENGUA
Objetivos	Contrastar ambas gramáticas	Analizar los errores de los estudiantes	Analizar la interlengua
Foco	La interferencia	Los errores	Toda la producción del aprendiente
Participación del alumno	nula	activa	activa
Valor del error	negativo	necesario	positivo
Causas del error	Interferencia con la L1	Interferencia con la L1, características de la L2	Indicadores de que el proceso de aprendizaje se está llevando a cabo.

Tabla Nº 1: Modelos de análisis

### 2.4. GLOSARIO DE TÉRMINOS

#### IL (interlengua)

**(LM) La lengua materna** es la que se adquiere cuando se es pequeño en un contexto natural, gracias a la interacción con la familia y la sociedad.

**lengua segunda (L2)** es la que se adquiere en contextos de inmersión en el país donde se habla esa lengua.

**lengua extranjera (LE)** se aprende en un país donde dicha lengua no es utilizada por los nativos en la vida cotidiana.

**lengua meta** es la lengua que se aprende o que se adquiere; con este término se suele hacer referencia tanto a una lengua segunda como a una lengua extranjera.

**Transferencia** es el grado de influencia de la gramática de la L1 respecto de la adquisición de L2. Este concepto está vinculado a la influencia de la lengua materna en la adquisición de la segunda lengua. Para la corriente del Análisis contrastivo, los hábitos adquiridos en el aprendizaje de la lengua materna son trasladados en el aprendizaje de otra lengua y se constituyen como la causa de los errores cometidos por los aprendices cuando no coinciden las estructuras de la L1 y la lengua meta.

Desde la teoría de la interlengua la transferencia responde a un mecanismo psicológico del sujeto que aprende y es uno de los elementos que ocasiona la fosilización.

**Interferencia** es entendida como el efecto de la lengua nativa sobre la lengua meta en el proceso de aprendizaje". (Gargallos pág. 37).

Se dice que hay interferencia cuando aparece algún rasgo fonético, morfológico, sintáctico o léxico característico de la lengua madre o nativa del aprendiz.

**Hipótesis acerca del acceso a la GU** ¿Por qué el estado final alcanzado en la adquisición de L2 es distinto al de la lengua materna?

Las hipótesis que apunta Rivero (2005, 25) son:

- Que la GU ya no esté disponible o si lo está dicha disponibilidad sea distinta.
- Que el input (los datos que recibe el aprendiente) tenga características distintas al input en L1 y que el modo de procesarlo sea diferente.
- Las diferencias no se hallarían en el conocimiento del lenguaje sino en el uso que se hace de ese conocimiento en L2
- Los factores que influyen sobre la producción son de acción radicalmente diferente y es lo que determina la variación.
- El conocimiento lingüístico no es equivalente al de la L1.

Esta cuestión hace referencia a la caracterización del Estado inicial de la L2. Este es un problema que da inicio a una serie de interrogantes que tienen qué ver con la disponibilidad de la GU y cómo accedemos a ella, para adquirir una segunda lengua.

Si se considera que el acceso a la GU es Total, es decir que ésta permanece abierta, disponible, habría que investigar qué sucede con la acción de la gramática universal en la L1: ¿Los parámetros que han sido fijados pueden resetearse, tienen una doble especificación?, ¿Cómo se elige el valor correspondiente a la nueva lengua?

En el caso de que la GU no se encontrara disponible, sino cerrada para adquirir otra lengua, ¿A partir de qué dispositivo se llevaría a cabo la adquisición?

Estas hipótesis pueden resumirse en tres siguiendo a White (2003) en Rivero (2006:30-40)

**1. Adquisición de L2 sin la GU** (Transferencia sin acceso): *el estado inicial en la adquisición de la lengua meta es la L1.* Página | 27

Los que sostienen esta hipótesis (Bley-Vroman 1990), argumentan que la GU se cierra o atrofia con la edad y por eso no sería accesible para aprender una segunda lengua. Por lo tanto se emplearían otros mecanismos cognitivos ya maduros para aprender una L2.

**2. Hipótesis de la transferencia completa:** *Adquisición de la L2 con acceso completo a la GU y con una gramática como estado inicial.* (Schwartz & Sprouse 1996)

El estado inicial difiere con respecto a la adquisición de la lengua materna ya que en la adquisición de L2 el estado inicial es la gramática de la L1. Sin embargo, el proceso de adquisición sería similar si se tiene en cuenta la participación de la GU, ya que se cree que las propiedades no instanciadas en la L1 aún están disponibles.

**3. Hipótesis de acceso completo sin transferencia** *GU como estado inicial.*

Proponen que la adquisición de la gramática de la L2 se realiza a partir de la interacción directa con los principios y parámetros de la GU con el input de la L2. (Epstein, Flynn & Martohardjono 1998)

El estado final alcanzado sería similar al de la L1 y la divergencia observada se explicaría como una cuestión del uso de ese conocimiento, no en términos de competencia.

## CAPÍTULO 3: EL CORPUS

Este capítulo explica cómo está organizado el corpus y la importancia que tiene la descripción de los datos en el ámbito investigativo de la adquisición de una segunda lengua. Describe el origen de los sujetos que forman parte de la prueba y muestra los tipos de consignas y ejercicios que han servido para la recolección de las producciones escritas.

### 3.1. Organización del corpus

El corpus empleado en esta investigación es de tipo textual, reúne textos escritos que han sido digitalizados y está dividido en cinco grupos según la lengua de origen de los aprendientes de español. Todas las producciones pertenecen a estudiantes de un nivel inicial. Así tenemos:

Grupo ALEMÁN (cuenta con 2488 palabras y 21 textos)

Grupo FRANCÉS (cuenta con 6796 palabras y 48 textos)

Grupo PORTUGUÉS (cuenta con 1721 palabras y 27 textos)

Grupo INGLÉS (cuenta con 1398 palabras y 30 textos)

Grupo HOLANDÉS (cuenta con 1600 palabras y 7 textos)

El criterio de clasificación se debe a la hipótesis lingüística inicial de la investigación, la cual propone que la causa de las desviaciones halladas en la adquisición de la flexión verbal irregular del español como L2 es intralingüística, obedeciendo a una particularidad de la lengua meta, sin importar qué lengua de origen posee el aprendiente.

A su vez, cada grupo está compuesto por una cantidad determinada de textos que han sido numerados y corresponden a producciones de diferentes sujetos. El objetivo es

facilitar el empleo de tablas en donde conste tipo de desviación, corpus y número de texto dentro del corpus.

El capítulo número 7 en donde se realiza el análisis estadístico responde entonces a un interrogante metodológico sobre el modo de organizar el corpus. El objetivo propuesto es analizar cada grupo de textos para extraer conclusiones en cuanto al porcentaje de formas verbales desviadas y de formas coincidentes con el español estándar, para determinar si existen o no diferencias significativas entre uno y otro grupo, con el fin de verificar la hipótesis ya mencionada.

### 3.2. Los sujetos

La presente investigación se basa en primer término en la preparación en sus aspectos lingüísticos e informáticos del corpus con el que se va a trabajar y en segundo término, en la selección de sujetos, estableciendo una muestra aleatoria de estudiantes extranjeros de español como L2, con no más de un año de estudio, pertenecientes a instituciones educativas de la Universidad de Rosario.

Se considera que el diseño muestral más apropiado para esta investigación es el estratificado, considerando como estratos a los grupos definidos por la lengua nativa. Luego se definirá el tamaño de la muestra y la asignación a cada estrato de acuerdo a las variables que se relevarán.

El corpus está formado por producciones de aprendientes de español, la franja etaria va desde los 16 años a los 45 años, predominando sujetos jóvenes de entre 17 y 25 años.

Estos sujetos se encuentran en Argentina con el propósito de estudiar una carrera universitaria en la Universidad de Rosario, de realizar sus estudios de Doctorado en la misma casa de estudios y de realizar un intercambio en localidades de Argentina.(grupo alemán)

La facultad de Humanidades y Artes recibe dos veces al año a cohortes intensivas de estudiantes extranjeros que vienen a cursar el Doctorado con mención en Ciencias de la Educación. Como profesora del cuerpo de docentes que dicta el seminario de “Español” para el Doctorado destinado a brasileños pude tomar pruebas y recoger el corpus perteneciente a la lengua de origen portugués. Es importante señalar que este corpus se encuentra en continuo crecimiento.

Otro grupo de producciones pertenecen a jóvenes que desean ingresar como estudiantes a la Universidad Nacional de Rosario, y son precisamente parte de esos exámenes de ingreso los que reuní como corpus, éstos corresponden a los grupos que tienen como lengua de origen al francés e inglés.

Para la selección de sujetos se eligieron a estudiantes que se hallan en una etapa inicial con el propósito de investigar qué sucede en cuanto a la adquisición de la morfología del español, teniendo como hipótesis de partida la manifestación de formas idiosincrásicas propias de la interlengua, que no se corresponden con las de la lengua madre ni con las de la lengua meta.

Los sujetos entonces se encuentran aprendiendo español y residen en Argentina, por eso prefiero denominar al aprendizaje del español como adquisición de segunda lengua.

Sin embargo, la mayoría de los sujetos que tiene como lengua materna al portugués sólo permanecen en Rosario cuatro semanas al año, luego no tienen contacto con la lengua española más allá de las clases tomadas en Brasil y éste es el segundo viaje que realizan.

Por lo tanto, poseen un nivel de español básico un tanto inferior al resto de los sujetos que conforman los otros corpus, que se encuentran viviendo en Argentina.

Según el marco de Referencia europeo<sup>11</sup>, el nivel corresponde a un usuario básico. Este primer nivel está constituido por dos subniveles: el A1 y el A2. Los aprendientes pertenecientes al corpus alemán, francés, holandés e inglés se encuentran en el subnivel A2, mientras que los aprendientes de origen portugués recién están alcanzando el A1.

Para aclarar esta diferencia veamos cuáles son los conocimientos mínimos que deben poseer los estudiantes encuadrados en este estrato:

A1 comprensión de lectura: Comprende palabras y nombres conocidos y frases muy sencillas; por ejemplo las que hay en letreros, carteles y catálogos.

A2 comprensión de lectura: Es capaz de leer textos muy breves y sencillos, encontrar información específica y predecible en escritos sencillos y cotidianos y comprende cartas personales breves y sencillas.

A1 interacción escrita: Puede escribir postales cortas y sencillas; por ejemplo, para enviar felicitaciones. Sabe rellenar formularios con datos personales; por ejemplo, su nombre, nacionalidad y dirección en el formulario del registro de un hotel. Dispone de un

---

<sup>11</sup> Disponible en [http://cvc.cervantes.es/ensenanza/biblioteca\\_ele/marco/cvc\\_mer.pdf](http://cvc.cervantes.es/ensenanza/biblioteca_ele/marco/cvc_mer.pdf)

repertorio básico de palabras y frases sencillas relativas a sus datos personales y a situaciones concretas.

A2 interacción escrita: Puede escribir notas y mensajes breves y sencillos relativos a sus necesidades inmediatas. Escribe cartas personales muy sencillas; por ejemplo, agradeciendo algo a alguien. Utiliza algunas estructuras sencillas correctamente, pero todavía comete, sistemáticamente, errores básicos.

Es importante aclarar que se eligieron sujetos con el mismo nivel de aprendizaje teniendo en cuenta que la investigación se basa en la hipótesis de que a pesar de trabajar con producciones escritas de aprendientes con diversas lenguas de origen no encontraremos diferencias en cuanto a la cantidad y al tipo de desviaciones y de formas coincidentes con el español estándar en lo que respecta a la flexión de verbos irregulares.

### 3.2.1. ¿Por qué “Segunda Lengua”?

Es necesario explicar a qué nos referimos con el término segunda lengua y distinguirlo de otros empleados en el campo de la adquisición, como por ejemplo el de “Lengua Extranjera”. La diferencia radica en las condiciones de desarrollo de la adquisición de la nueva lengua, es decir si esa lengua es o no la L1 de la comunidad en la que está inmerso el aprendiz que la toma como L2.

Como expone Rivero S. (2006) “La L2 es la L1 en un contexto en el que ésta se utiliza predominante y habitualmente para la comunicación...” (Rivero: 2006, 14).

En cambio, según algunas líneas teóricas, cuando hablamos de lengua extranjera, la L2 no es la lengua de comunicación habitual y por lo tanto implica una exposición parcial además de motivaciones diferentes a la hora de aprender la lengua. Por ejemplo, un sujeto que se va a vivir a otro país, el cual posee una lengua diferente necesita aprenderla para comunicarse diariamente, a diferencia de un sujeto que la adquiere en un contexto académico, a partir de un curso o por razones laborales.

Por lo ya dicho prefiero utilizar el término segunda lengua (L2)

Además se empleará “aprendiente”, “aprendiz” y “estudiante” como sinónimos para hacer referencia a los sujetos que forman parte de la investigación.

Siguiendo a Torruella- Llisterri (1999) la función principal de un corpus es la de establecer la relación entre la teoría y los datos y la efectividad en la preparación de un corpus reside en el hecho de poder mostrar una parte de la realidad lingüística que pretende estudiarse y para eso debe cumplir con dos características: primero estar organizado y segundo ser neutro, en el sentido de no poseer marcaciones para que nos permita recoger muestras proporcionales de todos los aspectos o niveles que quieran analizarse.

Cabe aclarar que la selección que se realiza al conformar un corpus nunca está libre de una determinación, es decir que planeamos un corpus desde un pensamiento o punto de vista determinado.

Los corpus informatizados representan una herramienta sumamente importante en el ámbito de la investigación lingüística ya que no sólo permiten ahorrar tiempo si no también trabajar de manera más ordenada y exhaustiva, sobre todo en lo que refiere al análisis gramatical automático o a traducción automática.

Los corpus son útiles en el campo de la confección de herramientas lingüísticas, en cuanto a la creación de diccionarios y también en el campo de la enseñanza de una segunda lengua,

Es importante señalar que la diferencia entre una recopilación de textos y un corpus es que éste último responde a un criterio interno que refiere a patrones lingüísticos presentes en los textos.

Torruella (1999) propone una clasificación de corpus según diversos criterios generales. El criterio que le corresponde al presente corpus es, según el porcentaje y la distribución de cada tipo de texto que lo compone, teniendo en cuenta que nos importa estudiar los escritos de aprendientes que tienen como lengua materna a diferentes lenguas.

Y dentro de las clases de corpus representa un corpus **comparable**:

“...son corpus que seleccionan textos parecidos en cuanto a sus características en más de una lengua o en más de una variedad” (1999:11), ya que la finalidad es comparar el



comportamiento de los escritos por aprendientes de diferentes lenguas en situaciones comunicativas parecidas.

Además se incluye en el tipo de corpus especializado ya que recoge textos que pueden aportar datos para la descripción de un tipo particular de lengua.

### 3.4. Los datos

En este apartado se explicará cómo se llevó a cabo la recolección de los datos. Tomando en cuenta las categorías que expone White (2003), tomado de Rivero (2006) sobre el tipo de datos requeridos en Adquisición de Segunda Lengua tenemos:

1. Datos de Producción: como la producción espontánea y elicitada.
2. Datos de Comprensión, como son los obtenidos a partir de tareas de identificación de imágenes
3. Datos de Intuición: como los juicios de gramaticalidad y los de valor de verdad.

En esta investigación se emplean datos de producción espontánea. Es necesario reparar en el hecho de que la ausencia de datos determinados de la producción espontánea no implica que estén restringidos, es decir que no formen parte de la competencia del estudiante.

Como el objetivo es describir y analizar los escritos mediante herramientas informáticas, no se considerarán para el trabajo computacional las emisiones o ausencias.

Por lo tanto, para achicar las limitaciones, propias de la investigación en el campo de adquisición se recolecta una cantidad considerable de datos que asegure el procesamiento estadístico y se delimita una muestra amplia y representativa.

Con respecto a los estímulos, tratan de controlarse, sobre todo en lo que respecta a las consignas para la producción que son dirigidas a los sujetos para no introducir variables intrusivas.

Los escritos responden a consignas sugeridas por el docente, por ejemplo: redactar un mail para reservar habitación de un hotel, escribir una carta a un amigo contándole lo que hiciste en tu estadía en la ciudad, comparar la ciudad de Rosario con su ciudad natal, completar con verbos, escribir en qué gustos coincide y en cuáles no con el personaje, entre otros.

A continuación se muestran algunas de las consignas propuestas para las producciones escritas y los textos de los estudiantes a modo de ejemplo:

**Enviar un e mail a un amigo para enviarle fotos que tomó durante la estadía en Rosario y para reclamarle el envío de las fotos tomadas por otra persona. (50 palabras)**

He venido a Rosario el mes pasado. Cuándo hemos hacido la fiesta el 10 de junio, vos haces una fotografia de nosotros. Es possible de me le enviar? Quiero tambien que evía las fotos para todas las personas del grupo. Muchas gracias para todas que hacía para nosotros. Me gusta mucho Rosario y me gusta mucho todo que hacemos en nos viaje

**Enviar un email al hotel solicitando información para realizar una reserva**

*¡Hola!*

*Somos dos canadienses que estarán de paso en la ciudad de Rosario el último fin de semana de septiembre. Necesitamos informacións sobre el hotel. ¿Cuál es la tarifa para dos días? Necesitamos también de saber si hay una habitación libre on el último fin de semana de septiembre. ¿Cuál son los modos de reserva? Nosotros somos mochos interesantes de vacaciones en la Argentina 4 esperamos ir a su hotel.*

*Muchas gracias.*

**¿En qué gustos no coincide con Gaby?**

*Podría decir que mis gustos son similar de los gustos de Gaby pero no me gusta el tenis. Soy un fanática del fútbol tambien. Me gusta tomer helados y me gusta passar tiempo con mi amigas. Pero no me gusta la tecnologia y las paginas web.*

*Gaby se gusta hacer páginas web. Me gusta helado también pero no comeo helados porque voy a "engordar".*

**Cuente su estadía en la ciudad de Rosario**

*Cuando llege en Rosario, 16 de enero de 2012, me sentí sorprendido con las novedades. El tiempo estaba muy agradable. Muchas cosas eran diferentes de mi ciudad. Había muchos perros por las calles de la ciudad. También había muchas parejas de viejos por toda la ciudad. Sentí muchas diferencias y las personas hablaban muy poco. Ellos casi no sonríen, parecían no gustar mucho de los brasileños. Ahora, julio de 2012, por segunda ocasión, me sentí más sereno. Vi muy pocos perros por las calles. La ciudad parecía mucho más bella pero el tiempo muy frío. Los profesores parecían más satisfechos. También el edificio de la universidad está más cómodo. La ciudad parece que está con mucho más personas. Cada vez estoy sintiéndome como un rosario. Estoy sintiéndome en mi propia casa. Cuando cese mi doctorado, quiero volver otras veces a Rosario.*

**Describe la escena retratada por el pintor Molina Campos respondiendo las siguientes preguntas:**

**¿Qué es? ¿Cómo es? ¿Cómo están?**

*El artista retrata la imagen de una casa de campo en la cual una familia humilde se agrupa en torno a una hoguera para calentarse del frío. Algunas figuras retratadas tales como: cajones, escoba, y pailas fijadas en las paredes, son muy interesantes, pues demuestran la simplicidad de la vida campestre incluyendo a las personas. El paisaje es agradable y muy común en la realidad de países como Brasil y Argentina.*

En este capítulo se describen las irregularidades que presenta el sistema verbal español, distinguiendo los cambios vocálicos de los cambios consonánticos y designando a un grupo de verbos como de irregularidades especiales. Este estudio servirá para la clasificación de los verbos irregulares que se encuentran en los textos que componen el corpus de interlengua.

### 4.1. ¿Cuándo un verbo se considera irregular?

Para el estudio de los verbos irregulares del español se recurre a la investigación realizada por Alcoba (2009) en la Gramática Descriptiva de la lengua española y se consulta el estudio emprendido por la Gramática de la Real Academia Española (2010).

Alcoba (1999) define a los verbos irregulares como aquellos verbos en los que en algunas de sus formas de flexión se altera parcialmente el constituyente radical (como en *cont- ar* *cuent- en* o en *pon-er* *pong- an*), cuando se altera la forma correspondiente del tema por supresión de la vocal temática (*pon* [e] (*d*) *ré*) o cuando los verbos emplean diferentes raíces que corresponden a orígenes etimológicos distintos como en los verbos *ser* o *ir*.

La caracterización que realiza de las irregularidades atañe a dos cuestiones: por un lado la manifestación fonológica de la irregularidad, ya sea por alternancia o incremento de segmentos, tanto de orden vocálico como consonántico y por otro lado, a la extensión de la irregularidad, referida tanto a las formas flexivas afectadas como a la cantidad de verbos a los que concierne dicha irregularidad.

No contempla a las variaciones ortográficas como manifestaciones de irregularidad flexiva y tampoco a las que provienen de alteraciones dadas por procesos fonológicos regulares en español como la alternancia vocal/semivocal en algunos verbos vocálicos, como ocurre en los verbos: *leer*, *proveer*, en donde se consonantiza en [y] la semivocal [j].<sup>12</sup>

---

<sup>12</sup> Siguiendo el criterio de Alcoba (1999) no incluí en este estudio a las variaciones que implican cambios de acentuación en los denominados verbos vocálicos, aquellos cuya raíz termina en vocal como por ejemplo el verbo *enviar*.

La RAE considera irregulares a los verbos que no se ajustan a los tres paradigmas de verbos regulares y también a los verbos vocálicos (que son los que tienen raíces que terminan en vocal como *leer, oír*), ya que en ellos se produce cambio de acento o cambios en el segmento vocálico (*leyó*).

No reconoce como irregularidades morfológicas a las variantes ortográficas que responden a ciertas reglas del sistema de la lengua española, como por ejemplo: *dirigimos, dirijamos*.

#### 4.1.1 ¿Cómo se manifiestan las irregularidades?

Tradicionalmente los verbos irregulares se clasificaban en cuatro grupos:

- a) Los que presentan irregularidades en la raíz de forma parcial.
- b) Las que se producen por supresión de la vocal temática.
- c) En las que se altera la raíz totalmente.
- d) Irregularidades ortográficas.

Una diferencia a destacar es que el último grupo de verbos no es incluido dentro del grupo de verbos irregulares por las gramáticas actuales.

La RAE agrupa las irregularidades en tres clases: vocálicas, cuando hay alternancia entre vocales (*pedir-pido*) o entre vocales y diptongos: (*contar-cuento*), consonánticas cuando se sustituye una consonante por otra (*hacer-hago*) y mixtas cuando se sustituyen vocales y consonantes simultáneamente como en *decir-digo*, aquí incluye a los verbos de pretéritos fuertes (*hizo-quiso-supó-pudo*).

Además de estas clases debemos mencionar que existen raíces supletivas (cuando el verbo presenta dos o más raíces distintas, por ejemplo *ir (voy)* o *ser (fui)* y verbos defectivos, los cuales no presentan formas que se desvíen de la conjugación regular sino que padecen la ausencia de determinadas formas en su paradigma de flexión. Por ejemplo: *abolir*.

Con respecto a la extensión de la irregularidad podemos distinguir:

- 1- La extensión externa, referida a la cantidad de verbos que afecta, por ejemplo, la irregularidad vocálica (e/i) involucra a los verbos terminados en (-entir, -erir, -ertir) como consentir, conferir, convertir.
- 2- La extensión interna indicando al conjunto de formas flexivas de un verbo a las que afecta la irregularidad. Así tenemos irregularidades que afectan a formas del tema de presente (que abarcan el presente del Modo Indicativo, el presente del Modo Subjuntivo y el Modo Imperativo), del tema de pretérito (que afecta al Pretérito del Indicativo, al Imperfecto y Futuro del Modo Subjuntivo) y del tema de futuro (que comprende Futuro del Indicativo y Condicional simple).

#### 4.2. Cambios vocálicos en la raíz

Las irregularidades vocálicas afectan a la vocal tónica de la raíz, y según la naturaleza de esta última pueden ser de orden palatal [e,i] como *sentir*, *sonreír* o de orden velar [o,u] como *dormir*, *volver*, *morir*.

Las alternancias se dan en las raíces tónicas del tema de presente:

- a) /e/ alterna en /ie/: empezar /empiezo
- b) /o/ alterna en /ue/: contar /cuento
- c) /i/ alterna en /ie/: adquirir/adquiero
- d) /u/ en /ue/: jugar /juego
- e) /e/ en /i/: pedir/pido

Como vemos, los cambios vocálicos pueden darse por diptongación o por cierre de vocal. Algunos verbos presentan simultáneamente más de una alternancia.

Dentro de este grupo se incluye a la alternancia que afecta a los verbos terminados en \_eir como freír y reír. (ríó-reí-rieron). Alcoba explica esta irregularidad como producto de la asimilación de la semivocal /i/. Dichos verbos son irregulares a causa del segmento [i] que provoca la manifestación alternante irregular de la /e/ radical.

#### 4.2.1 Cambios por diptongación

- a) Se produce la **alternancia e/ie** en todas las formas del singular y en la tercera persona del plural del presente del modo indicativo y del presente del subjuntivo. Afecta a verbos de la primera conjugación (acertar), de la segunda (entender) y de la tercera (sentir).

Página | 39

Algunos verbos de la primera conjugación que presentan la misma alternancia son:

Despertar, empezar, recomendar, pensar, cerrar.

- b) La alternancia **/o/ /ue/** también se da con verbos de las tres conjugaciones (contar- cuento, mover-muevo, dormir-duermo). Algunos otros verbos de la primera conjugación que presentan la misma irregularidad son: volar, probar, mostrar, almorzar, soñar, encontrar.

Con respecto a la segunda conjugación, los verbos terminados en **-ocer**, **oler**, **order** y **-over** presentan la misma alternancia. Por ejemplo: **cocer**, **doler**, **morder**, **envolver**.

- c) La alternancia **i- /ie/** se da en las mismas personas y tiempos verbales afecta sólo a verbos de la tercera conjugación como **adquirir** que alterna en **adquiero**.
- d) La irregularidad de alternancia **u /ue/** se presenta sólo en las formas de raíz tónica del presente del verbo **jugar** (juego).

#### 4.2.2. Cambios por cierre de vocal

- e) La irregularidad de cierre vocálico **e /i/** (pedir-pido) afecta a las formas fuertes del presente de indicativo y del subjuntivo y atañe a verbos de la tercera conjugación terminados en **edir** (medir), **-egir** (corregir), **-eguir** (perseguir), **-eñir** (reñir), **-estir** (vestir), **-etir** (derretir).

Además se extiende a todos los verbos terminados en **-eír**, por ejemplo: **freír** y **sonreír**.

Hay verbos que tienen la doble alternancia e /ie/ y cierre vocálico e/i como el verbo sentir (siento- sintió). A este mismo tipo de verbos pertenecen los verbos terminados en –entir (mentir: miento/mintió), –erir (herir: hiero/hirió), \_ertir (advertir: advierto/advirtió).

#### 4.3. Cambios consonánticos

Éstos pueden ser o bien por alternancia entre uno de los segmentos (hacer/hago) o bien por incremento de la consonante velar sorda [k] tras [Ø] en verbos como *conducir que (conduzco)*, o de la velar sonora [g] en verbos cuya raíz termina en [Ø, s, n, l] como *poner (pongo)*, *venir (vengo)*, o cuando la raíz termina en semivocal como *oír (oigo)*.

Dicho incremento de la consonante velar sonora /g/ se presenta también en verbos de raíz acabados en semivocal epéntica, como caer o traer (caigo-traigo). Presentan esta irregularidad en la primera persona singular del presente del modo indicativo (caigo) y en todo el tiempo presente del modo subjuntivo (caiga-caigamos).

Con respecto al verbo oír y los verbos acabados en –uir, Alcoba sostiene que interviene un proceso de consonantización de la semivocal (i) por epéntesis en el tema de presente. Argumenta que esta explicación regularizaría todos los verbos terminados en –uir ya que esta semivocal se consonantiza ante vocal (oyen-concluyen), se mantiene en la irregularidad de epéntesis en [g] (oigo, oigas) y se reduce ante /i/ tónica (oímos, concluimos)

#### 4.4 Irregularidades especiales

Se encuentran en este grupo los verbos: *haber, ser, estar, dar, ir*.

La particularidad de estos verbos es que se conjugan de manera exclusiva y hasta llegan a emplear dos raíces distintas como el caso de ser e ir.

El verbo estar presenta la peculiaridad de portar una raíz átona en todas sus formas.



En la tercera persona del singular del verbo haber usado como impersonal se manifiesta una semivocal (hay) y ocurre lo mismo en las primeras personas del singular del presente de los otros verbos del grupo: soy, estoy, doy, voy.

Exceptuando el verbo haber, que presenta otra irregularidad en el tema de futuro (habré) los demás verbos son regulares en las formas de futuro (seré-estaré-daré-iré).

#### 4.4.1. Pretéritos fuertes

Este tipo de verbos que son muy frecuentes en el uso, manifiestan el acento en la raíz de la 1era y 3era persona del singular, en las demás personas del tema de pretérito el acento recae en la vocal temática. Además, sin importar la conjugación del verbo la desinencia de 1era persona singular es /-e/ y de la tercera persona singular es /-o/, las otras desinencias son regulares. Presentan además variaciones vocálicas y consonánticas que escapan a una sistematización por su anomalía.

##### **Pretéritos fuertes con vocal a en la raíz:**

Traer (trajo)

##### **Pretéritos fuertes con vocal i en la raíz:**

Decir (dijo)

Hacer (hizo)

Querer (quiso)

venir (vino)

##### **Pretéritos fuertes con vocal u en la raíz:**

Andar (anduvo)

Caber (cupo)

Estar (estuvo)

Haber (hubo)

Poder (pudo)

Poner (puso)

Tener (tuvo)

La mayoría de estos verbos tienen alguna otra irregularidad en el tema de presente o en el tema de futuro.

En los verbos caber y saber, alterna además de manera exclusiva, la secuencia /ep/ de vocal consonante en el tema de presente: quepa/ quepamos, sepa/ sepamos y quepo y sé respectivamente para la primera persona del singular del modo Indicativo.

Las irregularidades que afectan al tema de futuro se caracterizan por la reducción de la vocal temática en verbos como saber, caber, haber, querer y poder (sabré-cabré. Habré-querré-podré).

En los verbos poner, tener, venir, valer, salir y sus derivados para formar el futuro se introduce el fonema /d/ (tendré- vendré-valdré- saldré)

#### 4.5. Consideraciones finales:

La irregularidad se puede manifestar en el incremento o alternancia de un segmento vocálico (poder/pude/puedo), en la alternancia o incremento de un segmento consonántico al final de la raíz, (parecer/parezco) por alternancias de los dos tipos en la raíz, (traer/traigo) en el uso de pretéritos particulares como el caso de los verbos de pretérito fuerte (estar/estuve), en alteraciones en las formas del futuro (poner/pondré), en el uso de participios especiales (morir/muerto) o en la combinación de dos de los fenómenos indicados.

Existe una extensión interna que refiere al conjunto de formas flexivas de un verbo a las que afecta la irregularidad y una extensión externa, referida a la cantidad de verbos que afecta.

## CAPÍTULO 5: LOS VERBOS IRREGULARES EN LA INTERLENGUA

En este capítulo se clasifican los verbos irregulares utilizados en la interlengua según la variación que presenten y se realiza además una descripción de los verbos hallados en cada corpus.

Página | 43

### 5.1. Clasificación de los verbos empleados en la interlengua

Aunque los verbos que presentan cambios ortográficos no son considerados irregulares por los lingüistas que se ocupan del tema, sino que son conceptualizados como verbos regulares que presentan variaciones ortográficas, son incluidos en esta investigación teniendo en cuenta que analizamos un corpus de textos escritos, en donde encontramos formas verbales pertenecientes a estos tipos de verbos, que presentan desviaciones respecto a las formas del español estándar. Por lo tanto, atendiendo al objetivo del trabajo, que es el de detectar estas formas verbales mediante los programas informáticos, es de utilidad que los analicemos.

Cabe aclarar que no se considerarán irregulares a los verbos de diptongo variable terminados en *-iar* (copiar) y *-uar* (actuar), dejando de lado las cuestiones referidas a acentuación, pero sí a los terminados en *-uir* como *construir* por la epéntesis de *-i-*: *construyo*.

Se toman los datos recogidos del corpus, considerando que los estudiantes se encuentran en un nivel inicial de la adquisición de segunda lengua y en la mayoría de los casos, sólo producen formas verbales en presente o en los pretéritos simples del modo indicativo, empleando la primera o la tercera persona del singular en gran medida.

En la interlengua encontramos formas verbales que corresponden a setenta verbos diferentes que sufren algún tipo de alteración o cambio.

A estos verbos vamos a describirlos según presenten irregularidad vocálica, irregularidad consonántica, cambios ortográficos o irregularidades especiales.

Por lo tanto, los verbos irregulares hallados en el corpus se clasifican según las alteraciones que los afectan:

A) Variaciones consonánticas, (tanto refiera a segmentos añadidos, como a consonantes que desaparecen o cambian por otras).

B) Variaciones vocálicas que se dan por diptongación ya sea de la vocal 'o' que alterna con el diptongo 'ue' o de la vocal 'e' que alterna con el diptongo 'ie'; o bien por un cierre de vocal (C) como ocurre en la alternancia *e / i*. Aquí también incluyo a los verbos terminados en \_eir que manifiestan la alternancia de la e en i.

D) Cambios gráficos que suceden cuando se da la presencia de consonantes como 'g' que hace "gu" si antecede a 'e' (agregar, jugar, pagar, rogar), 'c' que alterna con 'qu' (como en colocar, picar) o 'z' que cambia por 'c' (comenzar, localizar).

E) Por último, se ubican como verbos de irregularidades especiales a aquellos que presentan variaciones distintas. Son verbos como poder, traer, venir de pretérito fuerte. Ser, estar, dar e ir, que manifiestan la semivocal (y) y hacen soy, estoy, doy, voy. Otros como saber, haber, poder y querer, en los cuales hay una reducción de la vocal temática en tema de futuro, dando como resultado el grupo consonántico (sabré, podré). Verbos como tener, venir, que además de cambios vocálicos y consonánticos (tengo, tuve, vengo, vine), presentan cambios consonánticos en el futuro (pondrá, tendrá).

Irregularidad A: cambios consonánticos.

Irregularidad B: cambios en la raíz por diptongación.

Irregularidad C: cambios en la raíz por cierre de vocal.

Irregularidad D: cambios gráficos.

Irregularidad E: irregularidades especiales.

En la tabla que sigue se muestran los verbos irregulares con la indicación del grupo al que pertenecen:

Tabla de verbos irregulares de la interlengua

VERBOS	Consonánticos A	Diptonga B	Cierre C	Gráficos D	Especial E
agregar				X	
amenazar				X	
andar	X				
arrepentirse		X	X		
buscar				X	
calentar		X			
carecer	X				
colocar				X	
comenzar		X		X	
conducir	X				
conocer	X				
construir	X				
contar		X			
contribuir	X				
convertir		X	X		
costar		X			
dar					X
decir	X		X		
demostrar		X			
derretir			X		
desembarcar				X	
detener	X				X
dirigir				X	
doler		X			
empezar		X		X	
encontrar		X			
estar					X
freír			X		
haber	X				X
hacer	X				X
incluir					
ir	X				X
jugar		X		X	
localizar				X	
moler		X			
morir		X	X		
mostrar		X			
mover		X			

<b>nacer</b>	X				
<b>ofrecer</b>	X				
<b>oscurecer</b>	X				
<b>pagar</b>				X	
<b>parecer</b>	X				
<b>pedir</b>			X		
<b>pensar</b>		X			
<b>Picar</b>				X	
<b>poder</b>		X	X		X
<b>poner</b>	X		X		X
<b>preferir</b>		X	X		
<b>proponer</b>	X				X
<b>querer</b>	X	X	X		X
<b>recomendar</b>		X			
<b>recordar</b>		X			
<b>requerir</b>		X	X		
<b>renovar</b>		X			
<b>rogar</b>		X		X	
<b>saber</b>	X				X
<b>sacar</b>				X	
<b>salir</b>	X				
<b>sentarse</b>		X			
<b>sentir</b>		X	X		
<b>ser</b>					X
<b>sonreír</b>	X				
<b>sugerir</b>		X	X		
<b>suponer</b>	X				X
<b>tener</b>	X				X
<b>traer</b>	X				X
<b>ver</b> <sup>13</sup>					
<b>venir</b>	X		X		X
<b>volver</b>		X			

Tabla Nº 2: Verbos irregulares de la interlengua

El gráfico que sigue muestra la clasificación de irregularidades de los verbos de la interlengua:

<sup>13</sup> El verbo ver presenta la particularidad de agregar en algunas formss la vocal /e/ en Presente del Modo Indicativo y del Modo Subjuntivo (veo)

## VERBOS IRREGULARES DE LA INTERLENGUA

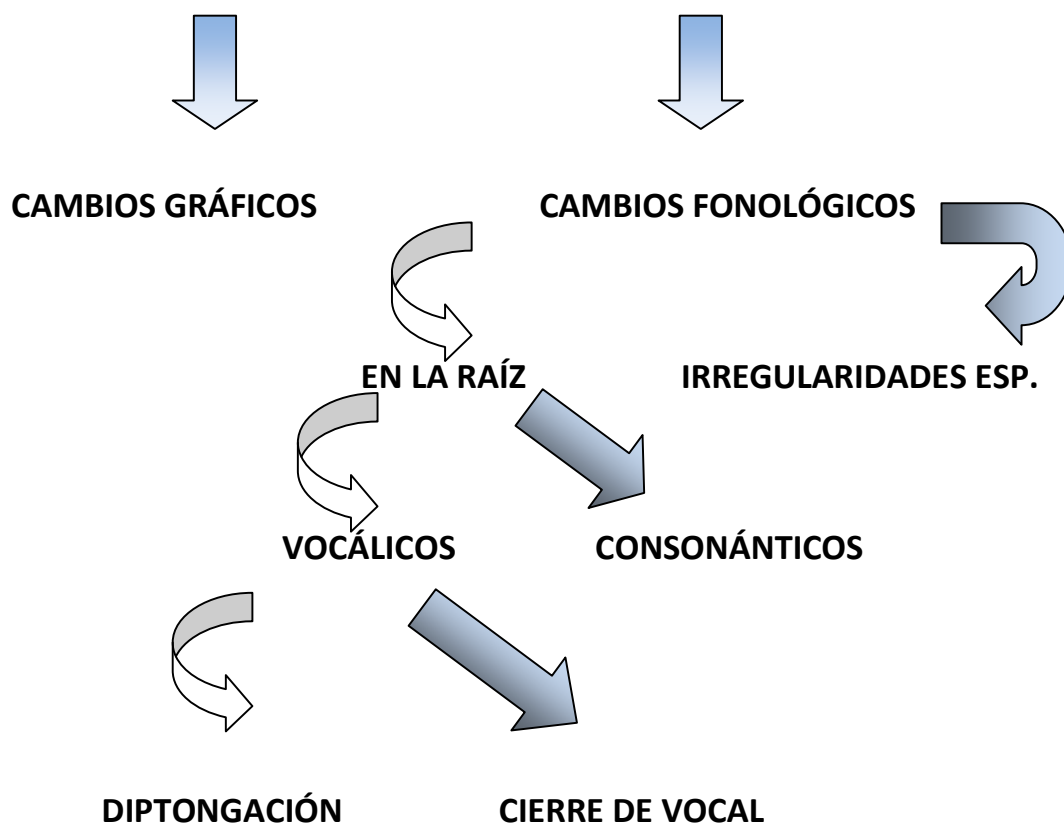


Gráfico Nº 2: Clasificación de irregularidades de la interlengua

En los cuadros que se presentan a continuación expongo los verbos irregulares de la interlengua agrupados según el tipo de irregularidad que presentan:

## CAMBIOS CONSONÁNTICOS

POR ALTERNANCIA	POR INCREMENTO
hacer	andar
decir	carecer
	carecer
	conducir
	conocer
	construir
	contribuir
	incluir
	nacer
	ofrecer
	oscurecer
	parecer
	poner
	proponer
	salir
	suponer
	tener
	traer
	venir

**Tabla Nº3: verbos con cambios consonánticos**

### Cambios vocálicos por diptongación

U en ue	O en ue	E en ie	I en ie
jugar	Volver	sugerir	
	Rogar	sentir	
	Renovar	sentarse	
	Recordar	requerir	
	Poder	recomendar	
	Contar	querer*	
	Costar	preferir	
	Demostrar	pensar	
	Doler	arrepentirse	
	Encontrar	calentar	
	Moler	comenzar	
	Morir	empezar	



Mover	pensar
-------	--------

**Tabla N° 4: verbos con cambios por diptongación**

**Cambios vocálicos por cierre de vocal**

E en i	O en u
pedir	morir
derretir	
freír	
pedir	

**Tabla N° 5: verbos con cierre de vocal**

**Cambios gráficos**

C por qu	G por Gu	Z por c	G por j
sacar	agregar	amenazar	dirigir
picar	pagar	localizar	
buscar			
colocar			
desembarcar			
picar			

**Tabla N° 6: Verbos con cambios gráficos**

**IRREGULARIDADES ESPECIALES**

Pretéritos fuertes	Verbos con dos raíces	Irregularidades especiales
andar	ser	dar
decir	ir	
detener		
estar		
haber		
hacer		
poner		

---

proponer (deriva de poner)

querer

saber

suponer (deriva de poner)

tener

traer

---

### **Tabla Nº7: verbos de irregularidades especiales**

\*verbo querer: además de esta variación (querer- quiero) presenta el pretérito fuerte (quiero- quise) y la síncope de la vocal temática en el futuro y condicional (querré-querría).

## **5.2. Verbos irregulares empleados según la L1**

### **En el corpus francés hallamos los siguientes verbos irregulares:**

AGREGAR, AMENAZAR, CALENTAR, COLOCAR, COMENZAR, CONDUCIR, CONOCER, CONSTRUIR, DAR, DECIR, DERRETIR, DESCUBRIR, DIRIGIR, EMPEZAR, ENCONTRAR, ES, ESTAR, FREIR, HABER, HACER, IR, MOSTRAR, NACER, PARECER, PEDIR, PENSAR, PICAR, PODER, PONER, QUERER, RECOMENDAR, RENOVAR, SACAR, SER, SUPONER, TENER, VENIR, VER, VOLVER.

### **En el corpus portugués encontramos los verbos que continúan:**

ANDAR, ARREPENTIRSE, CONOCER, COSTAR, DEMOSTRAR, DESEMBARCAR, EMBARCAR, EMPEZAR, ENCONTRAR, ESTAR, EXPONER, HABER, INCLUIR, IR, LLEGAR, OFRECER, PARECER, PENSAR, PROPONER, QUERER, REENCONTRARSE, SABER, SALIR, SENTIR, SER, SONREÍR, SUPONER, TENER, VENIR, VER.

### **En el corpus inglés localizamos los verbos que siguen:**

CONOCER, DAR, DORMIR, ESTAR, HABER, HACER, IR, PENSAR, PODER, PONER, PREFERIR, QUERER, SENTIR, SER, TENER, VENIR.

**En el corpus alemán tenemos los verbos:**

ANDAR, BUSCAR, CARECER, CONOCER, DAR, DECIR, DOLER, DORMIR, ELEGIR, EMPEZAR, ENCONTRAR, ESTAR, HABER, HACER, IR, JUGAR, PARECER, PEDIR, PODER, PONER, PREFERIR, PROPONER, QUERER, RECOMENDAR, RECORDAR, ROGAR, SABER, SACAR, SALIR, SENTARSE, SENTIR, SER, SOLER, SUGERIR, TENER, TRAER, VENIR, VER, VOLVER

**En el corpus holandés encontramos los siguientes:**

CONTAR, CONTRIBUIR, DEMOSTRAR, DETENER, ENCONTRAR, ESTAR, HABER, HACER, IR, LOCALIZAR, MOLER, MORIR, MOSTRAR, MOVER, NACER, OFRECER, PENSAR, PODER, QUERER, REQUERIR, SABER, SER, TENER, VOLVER.

Luego de analizar las formas verbales irregulares utilizadas por los aprendientes de las diferentes lenguas, se advierte que la desviación más frecuente es hacer regular un verbo que no lo es, por ejemplo: *sugirió* en vez de *sugirió*. En segundo lugar, observo que se asignan desinencias existentes en español pero no combinables con esas raíces, por ejemplo *queraba* por *quería* o no correspondiente: *haceo* por *hago*.

En muchos casos coinciden las dos desviaciones descriptas en una misma forma verbal: por ejemplo: *deteneron* por *detuvieron* o *habían hecho* en vez de *habían hecho*, o bien se presenta esta última con otro tipo de desviación<sup>14</sup>.

Otra alteración frecuente en la flexión de verbos irregulares es la de crear una forma verbal que presenta una variación distinta a la habitual, esto es cambio de vocal: *fuemos* (*fuimos*), cambio de consonantes: *trayera* (*trajera*) o posición incorrecta de diptongación como por ejemplo *pensio* por *pienso*.

Por último, en escasa cantidad se observan formas verbales que siguen la raíz correspondiente a las formas regulares del verbo y sin embargo son conjugadas como irregulares: *quierían* por *querían*.

Otro fenómeno observado que responde a una cuestión semántica pero que se detecta en la construcción de la forma verbal es el de condensar dos verbos de significado similar mezclando consonantes de ambos en una sola forma, por ejemplo: *pusque* en vez de *busque*, en donde la *p* proviene del verbo *poner*.

Cabe aclarar que en esta investigación, cuyo objetivo es el análisis automático de textos escritos, los verbos regulares que sufren variaciones gráficas son considerados dentro de la clasificación de verbos irregulares de la interlengua y por lo tanto las formas que

---

<sup>14</sup> Es importante mencionar que estos casos se contarán por dos en la tabla de la planilla Excel que se utilizará para comparar la cantidad de desviaciones de los corpus (ver capítulo N° 7).

refieren a la flexión de este tipo de verbos son tomadas como formas que se desvían de la lengua estándar.

En los casos mencionados observo que se mantiene la consonante del infinitivo (por ejemplo *coloce* (coloque), *pice* (pique), *agrega* (agregar) y no cambia por la correspondiente de igual sonido.

Por lo explicado anteriormente, encuadro las formas idiosincrásicas de la interlengua en cuatro casos desviados de la forma estándar y distinguidos con los números que van del 1 al 4, los cuales serán desarrollados en el apartado siguiente.

### 6.1. Clasificación de las formas verbales irregulares desviadas

**DESVIACIÓN 1:** Formas verbales irregulares que son conjugadas por los aprendientes como regulares. Por ejemplo: *queren* (quieren), *sacé* (saqué), *encontra* (encuentra), *volve* (vuelve), *empezo* (empiezo), *coloce* (coloqué), *renovan* (renuevan), *venieron* (vinieron).

**DESVIACIÓN 2:** Formas verbales regulares que son conjugadas por los aprendientes como irregulares. Por ejemplo: *quieremos* (queremos), *rueguía* (rogaba).

**DESVIACIÓN 3:** Formas verbales que presentan una irregularidad distinta a la estándar. Ejemplos: *demonstran* (demuestran), *he conoscado* (he conocido), *trayera* (trajera), *contriburion* (contribuyeron), *agaría* (haría), *pondiera* (pusiera).

**DESVIACIÓN 4:** Formas verbales que presentan una asignación incorrecta de desinencia, es decir, se construyen con terminaciones existentes en la lengua estándar pero que no corresponden a esas bases verbales, por ejemplo: *tenió* (tenía), *queraba* (quería), *dolo* (dolió), *ha estido* (ha estado), *sacieron* (sacaron).

## 6.2. Formas verbales irregulares desviadas por corpus

### 6.2.1. Corpus francés

En el corpus francés se hallan desviaciones del tipo 1, regularizar formas que son irregulares, en verbos que presentan cambios en la raíz por diptongación (*calente* por *caliente*, *empeza* por *empieza*, *renovan* por *renuevan*) y en verbos regulares con cambios ortográficos (*coloce* por *coloque*, *pice* por *pique*).

DESVIACIÓN 1	
COLOCE (coloque)	COLOCAR
PICE (pique)	PICAR
CALENTE (caliente)	CALENTAR
FREE (fría)	FREIR
SACE (saque)	SACAR
EMPEZA (empieza)	EMPEZAR
EMPEZO (empiezo)	EMPEZAR
RENOVAN (renuevan)	RENOVAR
COMENCO (comienzo)	COMENZAR

Tabla Nº 7 ejemplos de desv1 (francés)

En cuanto a desviaciones del tipo 2, con verbos que manifiestan cambios por diptongación (*volver*, *comenzar*).

DESVIACIÓN 2	
<b>VUELVÉ (volvió)</b>	VOLVER
<b>COMIENZAMOS (comenzamos)</b>	COMENZAR

Tabla N° 8: ejemplos de desv2 (francés)

Respecto del tipo de desviación 3, formas que presentan otro tipo de variación encontramos verbos con cambios gráficos que son conjugados con ausencia del grafema – u- (picar, agregar, colocar) además de otros.

DESVIACIÓN 3	
<b>PIQE (pique)</b>	PICAR
<b>AGREGE (agregue)</b>	AGREGAR
<b>COLOQE (coloque)</b>	COLOCAR
<b>CONCÍ (conocí)</b>	CONOCER

Tabla N°9: ejemplos de desv3 (francés)

Por último, en las formas desviadas del tipo 4, hallamos verbos, formados con desinencias distintas a las correspondientes, que también presentan otro tipo de desviación como por ejemplo hacer irregular la raíz: *vuelve* (por volvió).

DESVIACIÓN 4	
<b>FREE (fría )</b>	FREIR
<b>SACIERON (sacaron)</b>	SACAR
<b>TENÉ (tenía)</b>	TENER
<b>ESTUVA (estuve)</b>	ESTAR
<b>VIA (vio)</b>	VER
<b>VUELVÉ (volvió)</b>	VOLVER
<b>TENGES (tengas)</b>	TENER
<b>PUEDÁIS (puedes)</b>	PODER

Tabla N° 10: ejemplos de desv4 (francés)

### 6.2.2. Corpus inglés

Encontramos casos de desviaciones del tipo 1 en verbos que diptongan la raíz como querer y pensar y en verbos de irregularidades especiales como hacer y haber. Cabe aclarar que no se hallan verbos que presenten cambios de grafía.

Página | 56

DESVIACIÓN 1	
QUEREN (quieren)	QUERER
HACIDO (hecho)	HACER
HACIEMOS (hicimos)	HACER
HAMOS VISITADO (hemos visitado)	HABER
HAMOS APRENDIDO (hemos aprendido)	HABER
PENSIO (pienso)	PENSAR

Tabla N°11: ejemplos de desv1 (inglés)

Para el tipo de desviación 2 hallamos verbos que diptongan la raíz como querer y poder. Las formas idiosincrásicas analizadas tienen la particularidad de conjugar como irregular la primera y la tercera persona del plural, que precisamente son las formas que se mantienen regulares en los verbos que siguen el mismo modelo.

DESVIACIÓN 2	
QUIERÍAN (querían)	QUERER
QUIEREMOS (queremos)	QUERER
PUEDEMOS (podemos)	PODER

Tabla N° 12: ejemplos de desv2 (inglés)

En cuanto a desviaciones del tipo 3, éstas se dan en verbos de irregularidades especiales como tener, dar e ir. En el primer caso se repite la consonante nasal, en el segundo el verbo se mantiene en infinitivo y suma la desinencia adecuada, en el tercero no sólo regulariza la raíz sino que además agrega la vocal e, similar a lo observado en el ejemplo



siguiente: *pensio*. En la forma desviada he conoscado, se observa además de la colocación de una desinencia inadecuada la presencia de la consonante sibilar. En el último ejemplo confunde la primera persona (fui) con la tercera (fue) y forma *fuemos*.

DESVIACIÓN 3	
TIENNES (tienes)	TENER
DARES (des)	DAR
HACIEMOS (hicimos)	HACER
PENSIO (pienso)	PENSAR
HE CONOSCADO (he conocido)	CONOCER
FUEMOS (fuimos)	IR

Tabla Nº 13: ejemplos de desv3 (inglés)

En lo que respecta a las desviaciones del tipo 4, vemos que se intercambian las desinencias propias de participios (ado/ido).

DESVIACIÓN 4	
VEIS (vas)	IR
HE CONOSCADO (he conocido)	CONOCER
HA ESTIDO (ha estado)	ESTAR

Tabla Nº 14: ejemplos de desv4 (inglés)

### 6.2.3. Corpus portugués

Hallamos desviaciones 1 en verbos que diptongan la raíz (encontrar, pensar, tener), y en verbos de pretéritos fuertes como proponer y exponer que derivan de poner.

DESVIACIÓN 1	
HA PROPOSTO (propuesto)	PROPONER
EXPONISTE (expusiste)	EXPONER
PROPONÍ (propuse)	PROPONER
ENCONTRASE (se encuentra)	ENCONTRAR
ENCONTRAN (encuentran)	ENCONTRAR
ENCONTRA (encuentra)	ENCONTRAR
LLEGE (llegué)	LLEGAR
TENE (tiene)	TENER
PENSO (pienso)	PENSAR

Tabla Nº 15: ejemplos de desv1 (portugués)

Para las desviaciones del tipo 2 tenemos un solo caso en un verbo con cambio vocálico por diptongación.

DESVIACIÓN 2	
ENCUENTREI (encontré)	ENCONTRAR

Tabla Nº 16: ejemplos de desv2 (portugués)

En cuanto a desviaciones de tipo 3 los ejemplos exhiben una terminación inexistente en el español: –ei- (*desembarquei, lleguei, embarquei, estarei*). Hay verbos que diptongan en la raíz (*costar, demostrar*) que presentan diptongación de otro tipo (*cuosta*). También hay formas verbales en donde se intercambian consonantes (*incruyo*) o se agregan (*demonstré*).

DESVIACIÓN 3	
Forma desviada	Verbo irreg.
INCRUYO (incluyo)	INCLUIR
ARREPINTINRON(arrepintieron)	ARREPENTIRSE
TIEM (tiene)	TENER
HA (hay)	HABER
DEMONSTRAN (demuestran)	DEMOSTRAR

<b>INCLUJENDO (incluyendo)</b>	INCLUIR
<b>DEMONSTRÉ (demostré)</b>	DEMOSTRAR
<b>TIBE (tuve)</b>	TENER
<b>DESEMBARQUEI (desembarqué)</b>	DESEMBARCAR
<b>EMBARQUEI (embarqué)</b>	EMBARCAR
<b>LLEGUEI (llegué)</b>	LLEGAR
<b>ENCUENTREI (encontré)</b>	ENCONTRAR
<b>ESTAREI (estaré)</b>	ESTAR
<b>ESTIBE (estuve)</b>	ESTAR
<b>PARECIEN (parecen)</b>	PARECER
<b>HE EMPEZATO (empezado)</b>	EMPEZAR
<b>ESTIVE (estuve)</b>	ESTAR
<b>TIEN (tiene)</b>	TENER
<b>CUOSTA (cuesta)</b>	COSTAR

Tabla Nº17: ejemplos de desv3 (portugués)

Respecto de las desviaciones que presentan desinencias existentes pero no combinables con esas bases, encontramos verbos de irregularidades especiales como estar y proponer que es de pretérito fuerte.

<b>DESVIACIÓN 4</b>	
<b>Forma desviada</b>	Verbo irreg.
<b>ESTOE (estoy)</b>	ESTAR
<b>PROPONAN (proponen)</b>	PROPONER
<b>ESTAE (está)</b>	ESTAR

Tabla Nº 18: ejemplos de desv4 (portugués)

#### 6.2.4. Corpus alemán

En el corpus alemán, la mayoría de las desviaciones 1 provienen de verbos con cambios vocálicos tanto sea por cierre de vocal (pedir, venir, sentir) como por diptongación (sugerir, volver, jugar, preferir), también se da la regularización de formas irregulares en

verbos de pretérito fuerte como proponer y traer. En un solo caso encontramos un verbo con cambio gráfico (sacar).

DESVIACIÓN 1	
Forma desviada	Verbo irreg.
SUGERÍO (sugirió)	SUGERIR
PROPONÍ(propuse)	PROPONER
SACÉ (saqué)	SACAR
PEDO(pido)	PEDIR
PIDERON (pidieron)	PEDIR
TRAERON (trajeron)	TRAER
VENIERON (vinieron)	VENIR
HACISTE (hiciste)	HACER
PREFERIO (prefirió)	PREFERIR
VOLVE (vuelve)	VOLVER
JUGA (juega)	JUGAR
SENTIÓ (sintió)	SENTIR

Tabla Nº 19: ejemplos de desv1 (alemán)

En cuanto al segundo grupo de desviaciones obtenemos sólo un caso con un verbo que diptonga (*rueguía*).

DESVIACIÓN 2	
Forma desviada	Verbo irreg.
RUEGUÍA (rogaba)	ROGAR

Tabla Nº 20: ejemplos de desv2 (alemán)

Respecto de la desviación 3, tenemos formas idiosincrásicas a partir de la flexión de verbos que presentan irregularidades especiales como hacer (*agaría*), de pretérito fuerte como traer (*trayera, trayese*) y poner (*poniera*). En menor medida se dan en verbos con cambios gráficos (pagar, buscar).

DESVIACIÓN 3	
Forma desviada	Verbo irreg.
TRAYERA (trajera)	TRAER
TRAYESE (trajese)	TRAER
PONDIERA (pusiera)	PONER
PAGUERAN (pagaran)	PAGAR
AGARÍA (haría)	HACER
PUSQUE (busque)	BUSCAR
EQUENTRE (encuentre)	ENCONTRAR
PUSQUE (busque)	BUSCAR

Tabla Nº 21: ejemplos de desv3 (alemán)

Por último, para el caso de formas verbales que se conjugaban con una desinencia distinta, encontramos verbos de irregularidades especiales como *querer* (*queraba*), en donde se emplea la desinencia propia de los verbos de la primera conjugación en lugar de la desinencia correspondiente a un verbo de la segunda. El mismo procedimiento pero a la inversa se da en las formas desviadas *rueguía* y *pagueran*. En otros casos se observa el uso de desinencias propias del tiempo presente para verbos conjugados en pretérito: *estuben* y *dolo*.

DESVIACIÓN 4	
Forma desviada	Verbo irreg.
RUEGUÍA (rogaba)	ROGAR
PAGUERAN (pagaran)	PAGAR
ESTUBEN (estuve)	ESTAR
QUERABA (quería)	QUERER
DOLO (dolió)	DOLER

Tabla Nº 22: ejemplos de desv4 (alemán)

### 6.2.5. Corpus holandés

Es importante señalar que en este corpus no se hallan desviaciones del tipo 2.

Observamos que se regularizan verbos que diptongan la vocal o por ue como *moleron*, *conta*, *pode*, *encontra*. Verbos de pretérito fuerte como tener (*tenimos*, *tenió*) y su derivado detener (*deteneron*) y un solo caso de cambio gráfico (*contriburion* por *contribuyeron*).

DESVIACIÓN 1	
Forma desviada	Verbo irreg.
PODE (puede)	PODER
MOLERON (molieron)	MOLER
CONTA (cuenta)	CONTAR
DETENERON (detuvieron)	DETENER
ENCONTRA (encuentra)	ENCONTRAR
CONTRIBURION (contribuyeron)	CONTRIBUIR
TENIMOS (tuvimos)	TENER
TENIÓ (tuvo)	TENER

Tabla Nº 23: ejemplos de desv1 (holandés)

Encontramos desviaciones del tipo 3 en verbos que diptongan como demostrar, mostrar, y requerir que presentan distintas particularidades. Algunos son flexionados con el incremento de consonante como *demonstran*, en cambio en otros se cierra la vocal y no diptonga (*requiere*). En cuanto a verbos con cambios gráficos sólo hallamos localizar.

DESVIACIÓN 3	
Forma desviada	Verbo irreg.
HACEROS (hacen)	HACER
DEMONSTRAN (muestran)	MOSTRAR
DEMONSTRAN (demuestran)	DEMOSTRAR
DETENERON (detuvieron)	DETENER
REQUIRE (requiere)	REQUERIR

<b>LOCALICÓ (localizó)</b>	LOCALIZAR
<b>OFFRÓ (ofreció)</b>	OFRECER

Tabla N° 24: ejemplos de desv3 (holandés)

Para finalizar, en el último caso de desviaciones encontramos formas con otras desinencias al conjugar el verbo tener de pretérito fuerte.

<b>DESVIACIÓN 4</b>	
<b>Forma desviada</b>	Verbo irreg.
<b>TENIÓ (tuvo)</b>	TENER
<b>TENÍS (tenés)</b>	TENER

Tabla N° 25: ejemplos de desv4 (holandés)

### 6.3. Consideraciones finales

Del análisis descriptivo de las formas verbales irregulares desviadas concluyo que los verbos que diptongan en la raíz son los que presentan mayor cantidad de desviaciones del tipo 1, es decir, los aprendientes de español regularizan los verbos que diptongan a la hora de flexionarlos en presente, esto sucede con verbos de las tres conjugaciones como encontrar, pensar, calentar, renovar, jugar, mostrar, (para la primera), tener, querer, poder, volver (para la segunda) y sugerir y preferir (para la tercera conjugación.). En menor medida encuentro anomalías en verbos que sufren cambios ortográficos.

Las desviaciones del tipo 2 también se manifiestan en verbos de diptongación. Se emplean las raíces irregulares de estos verbos (por ejemplo querer, poder, rogar, mostrar) en la conjugación de personas verbales que no presentan esa variación como ocurre con la primera persona plural del presente y así obtenemos formas como *queremos*, *rueguía*, *puedemos*, entre otras.

Respecto a las desviaciones del tercer grupo, que escapan a una sistematicidad, remiten a verbos de irregularidades especiales como tener, ser, dar, estar y a verbos que diptongan en la raíz. En menor medida aparecen formas flexionadas de verbos con cambios consonánticos como conocer, parecer y de verbos con cambios gráficos como buscar y colocar.

En cuanto a las desviaciones del tipo 4, que se forman con desinencias del español pero que no corresponden a esas raíces considero que en su mayoría corresponden a verbos de irregularidades especiales como estar y tener.



## La lingüística computacional

Para introducir el análisis con herramientas informáticas explicaré brevemente a qué refiere el término lingüística computacional y cuáles son sus aplicaciones.<sup>15</sup>

Página | 65

En los últimos años se ha acuñado el término tecnologías del lenguaje para referirse a todas aquellas tareas en las que se aplica el conocimiento sobre la lengua para desarrollar sistemas informáticos capaces de reconocer, analizar, interpretar y generar lenguaje. Como explica Lavid (2005) existen tres procesos básicos de las tecnologías del lenguaje:

1. Introducción de datos de entrada en el sistema.
2. Tratamiento de los datos en diferentes niveles.
3. Producción de los datos de salida.

En los últimas tres décadas surge la Lingüística Computacional como ciencia del lenguaje que contribuye al conocimiento de los procesos cognitivos de comprensión y producción del lenguaje, combinando teorías del lenguaje natural con técnicas estadísticas.

*¿Cuál es la diferencia de la LC con la Lingüística descriptiva?*

La ventaja de utilizar modelos computacionales del lenguaje permite la comprobación de teorías lingüísticas ya que los procesos pueden inspeccionarse y experimentarse a través de la construcción de programas y bases de conocimiento.

---

<sup>15</sup> Ver Solana Z., Tramallino C., Méndez B. y otros (2014) *Estudios del lenguaje: niveles de representación lingüística* (Proyecto Latin. Iniciativa latinoamericana de escritura de libros abiertos).

### *¿Qué es la Lingüística Computacional?*

Es un área interdisciplinaria que toma saberes de la Lingüística y la Informática y su tarea consiste en crear sistemas informáticos capaces de procesar el lenguaje humano.

### *¿Qué objetivo se propone la LC?*

Diseñar programas informáticos que puedan emular en vez de simular la capacidad lingüística humana. Emular no significa comprender cómo funciona el cerebro humano si no intentar construir sistemas que comprendan y produzcan el lenguaje de manera similar a un humano.

### *¿Cuáles son las aplicaciones de la LC?*

- La comunicación entre el hombre y las máquinas
- La comunicación entre personas que hablan diferentes lenguas.
- Las tareas lingüísticas como el análisis de textos, la creación de diccionarios electrónicos o de correctores.
- El dominio de la Web.

## CAPÍTULO 7: COMPARACIÓN DE CORPUS DE DISTINTAS LENGUAS DE ORIGEN MEDIANTE TÉCNICAS ESTADÍSTICAS

El objetivo de este capítulo es el de emplear técnicas estadísticas y herramientas informáticas para determinar qué cantidad de formas verbales irregulares son coincidentes con la lengua estándar y qué cantidad presenta desviaciones respecto a ésta. Las conclusiones dependerán de la obtención de diferencias estadísticamente significativas entre los distintos corpus con los que se trabaja.

En cuanto a las herramientas informáticas, utilizaré la planilla Excel, la cual me permitirá realizar tablas dinámicas que servirán para el posterior trabajo estadístico para el que se empleará la técnica estadística: prueba no paramétrica de Kruskal- Wallis.

### 7.1. HERRAMIENTA INFORMÁTICA: Planilla Excel

Luego de examinar el corpus de interlengua y de extraer todas las formas verbales irregulares se confecciona una tabla por corpus según la lengua de origen de los aprendientes de español.

Una vez obtenidas las ocurrencias coincidentes con la lengua estándar y las ocurrencias desviadas, creamos en la planilla Excel una tabla, a la que llamaremos tabla 1. Ésta dispone una matriz de dimensión: tantas filas como cantidad de ocurrencias lingüísticas analicemos y tantas columnas como corpus +texto+ ocurrencia+ verbo+ desviaciones. En primer lugar analizamos el corpus correspondiente a la lengua francesa, ubicando en la columna siguiente el número de texto correspondiente a la ocurrencia, asignando una fila para cada una. Si hubiera ocurrencias repetidas se escriben cuantas veces hayan aparecido. En la cuarta columna se coloca el verbo irregular del que proviene y en la última el tipo de desviación que presenta, teniendo como posibilidades DESV1, DESV2, DESV3, DESV 4 o EST si coincide con la lengua meta. Con los demás corpus realizamos las mismas tablas, una a continuación de la otra.

Fragmento de la tabla 1 perteneciente al corpus francés<sup>16</sup>

CORPUS	TEXTO	OCURRENCIA	VERBO	TIPO DESV_
FRANCÉS	8	TENÍAN	TENER	EST
FRANCÉS	8	HIZO	HACER	EST
FRANCÉS	9	DERRITA	DERRETIR	EST
FRANCÉS	9	PONGA	PONER	EST
FRANCÉS	9	PONGA	PONER	EST
FRANCÉS	9	COLOCE	COLOCAR	DESV1
FRANCÉS	9	PICE	PICAR	DESV1
FRANCÉS	9	CALENTE	CALENTAR	DESV1
FRANCÉS	9	PIQE	PICAR	DESV3
FRANCÉS	9	FREE	FREIR	DESV1
FRANCÉS	9	FREE	FREIR	DESV4
FRANCÉS	9	AGREGE	AGREGAR	DESV3
FRANCÉS	9	COLOQE	COLOCAR	DESV3
FRANCÉS	10	ESTÁS	ESTAR	EST
FRANCÉS	10	ES	SER	EST
FRANCÉS	11	SACE	SACAR	DESV1
FRANCÉS	12	PUEDE	PODER	EST
FRANCÉS	12	PUEDO	PODER	EST
FRANCÉS	12	ESTÁN	ESTAR	EST
FRANCÉS	12	ESTOY	ESTAR	EST
FRANCÉS	13	FUERON	IR	EST
FRANCÉS	13	FUERON	IR	EST
FRANCÉS	14	ES	SER	EST
FRANCÉS	14	TENGO	TENER	EST
FRANCÉS	14	TIENE	TENER	EST
FRANCÉS	15	FUERON	IR	EST
FRANCÉS	15	SACIERON	SACAR	DESV4
FRANCÉS	15	CONCÍ	CONOCER	DESV3
FRANCÉS	15	FUERON	IR	EST

Tabla Nº 26: fragmento de Tabla dinámica 1 (francés)

<sup>16</sup> Ver tablas Nº 1 completas en la sección Anexos.

Fragmento de tabla N° 1 para el corpus alemán

CORPUS	TEXTO	OCURRENCIA VERBO		TIPO DESV_
ALEMÁN	1	SUGERÍÓ	SUGERIR	DESV1
ALEMÁN	1	RUEGUÍA	ROGAR	DESV2
ALEMÁN	1	RUEGUÍA	ROGAR	DESV4
ALEMÁN	1	TRAYERA	TRAER	DESV3
ALEMÁN	1	TRAYESE	TRAER	DESV3
ALEMÁN	1	PONDIERA	PONER	DESV3
ALEMÁN	1	PAGUERAN	PAGAR	DESV3
ALEMÁN	1	PAGUERAN	PAGAR	DESV4
ALEMÁN	1	PIDIERA	PEDIR	EST
ALEMÁN	1	DIJO	DECIR	EST
ALEMÁN	1	DIJE	DECIR	EST
ALEMÁN	1	QUERÍA	QUERER	EST
ALEMÁN	1	DIJO	DECIR	EST
ALEMÁN	1	PAGABAN	PAGAR	EST
ALEMÁN	1	DIJO	DECIR	EST
ALEMÁN	1	DIJE	DECIR	EST
ALEMÁN	1	PODÍAN	PODER	EST
ALEMÁN	1	DIJE	DECIR	EST
ALEMÁN	1	DIJE	DECIR	EST
ALEMÁN	1	TUVIERAN	TENER	EST
ALEMÁN	2	RECOMIENDA	RECOMENDAR	EST
ALEMÁN	2	SUGIERE	SUGERIR	EST
ALEMÁN	3	SABEN	SABER	EST
ALEMÁN	3	ESTOY	ESTAR	EST
ALEMÁN	3	TENGO	TENER	EST
ALEMÁN	3	QUIERO	QUERER	EST
ALEMÁN	3	DIJE	DECIR	EST
ALEMÁN	3	HAYA QUERIDO	QUERER	EST
ALEMÁN	3	DIJERON	DECIR	EST
ALEMÁN	3	RECOMENDARON	RECOMENDAR	EST
ALEMÁN	3	ELIGIERA	ELEGIR	EST
ALEMÁN	3	SÉ	SABER	EST
ALEMÁN	3	PROPONÍ	PROPONER	DESV1
ALEMÁN	3	SACÉ	SACAR	DESV1

ALEMÁN	3	AGARÍA	HACER	DESV3
ALEMÁN	4	TENÍA	TENER	EST
ALEMÁN	4	EMPEZAMOS	EMPEZAR	EST
ALEMÁN	4	TIENEN	TENER	EST
ALEMÁN	4	EMPEZÓ	EMPEZAR	EST
ALEMÁN	4	TRAJERON	TRAER	EST
ALEMÁN	4	ERA	SER	EST
ALEMÁN	4	DORMÍ	DORMIR	EST
ALEMÁN	4	HABÍA VISTO	VER	EST
ALEMÁN	4	FUIMOS	IR	EST
ALEMÁN	4	FUIMOS	IR	EST
ALEMÁN	4	FUE	IR	EST
ALEMÁN	4	PUDE	PODER	EST
ALEMÁN	4	SENTÍ	SENTIR	EST
ALEMÁN	4	SENTÉ	SENTARSE	EST

Tabla Nº 27: fragmento de Tabla dinámica 1 (alemán)

## Fragmento de la tabla Nº 1 correspondiente al corpus inglés

CORPUS	TEXTO	OCURRENCIA	VERBO	TIPO DESV_
INGLÉS	17	ES	SER	EST
INGLÉS	18	VEIS	IR	DESV4
INGLÉS	19	QUIERO	QUERER	EST
INGLÉS	20	TENGO	TENER	EST
INGLÉS	20	VAN	IR	EST
INGLÉS	21	QUIERO	QUERER	EST
INGLÉS	21	PUEDES	PODER	EST
INGLÉS	22	SOY	SER	EST
INGLÉS	22	QUIERO	QUERER	EST
INGLÉS	22	ES	SER	EST
INGLÉS	22	SON	SER	EST
INGLÉS	22	QUEREMOS	QUERER	EST
INGLÉS	22	SOMOS	SER	EST
INGLÉS	23	HACIDO	HACER	DESV1
INGLÉS	23	HACIEMOS	HACER	DESV1

INGLÉS	23	HACIEMOS	HACER	DESV3
INGLÉS	23	HACÉS	HACER	EST
INGLÉS	23	QUIERO	QUERER	EST
INGLÉS	23	QUEREMOS	QUERER	EST
INGLÉS	23	TENEMOS	TENER	EST
INGLÉS	23	SOMOS	SER	EST
INGLÉS	23	PUEDE	PODER	EST
INGLÉS	23	SOY	SER	EST
INGLÉS	23	SOY	SER	EST
INGLÉS	23	QUIERO	QUERER	EST
INGLÉS	23	PONES	PONER	EST
INGLÉS	23	QUIERO	QUERER	EST
INGLÉS	23	HE VENIDO	VENIR	EST
INGLÉS	24	QUEREMOS	QUERER	EST
INGLÉS	24	ES	SER	EST
INGLÉS	24	TENEMOS	TENER	EST
INGLÉS	25	PUEDAMOS	PODER	DESV1
INGLÉS	25	QUIEREMOS	QUERER	DESV2
INGLÉS	25	SOMOS	SER	EST
INGLÉS	25	ES	SER	EST
INGLÉS	25	PUEDE	PODER	EST
INGLÉS	26	QUIERO	QUERER	EST
INGLÉS	26	QUIERO	QUERER	EST
INGLÉS	26	PENSIO	PENSAR	DESV1
INGLÉS	26	PENSIO	PENSAR	DESV3
INGLÉS	27	ES	SER	EST
INGLÉS	27	ES	SER	EST
INGLÉS	27	VOY	IR	EST
INGLÉS	27	SON	SER	EST
INGLÉS	27	QUERÍA	QUERER	EST
INGLÉS	27	SON	SER	EST
INGLÉS	27	QUIERO	QUERER	EST
INGLÉS	27	ES	SER	EST
INGLÉS	27	ES	SER	EST
INGLÉS	27	QUIERO	QUERER	EST
INGLÉS	27	SON	SER	EST
INGLÉS	27	SERÍA	SER	EST
INGLÉS	27	ES	SER	EST
INGLÉS	27	DORMIMOS	DORMIR	EST

INGLÉS	27	HAMOS VISITADO	HABER	DESV1
INGLÉS	27	HAMOS APRENDIDO	HABER	DESV1
INGLÉS	27	HE CONOSCADO	CONOCER	DESV3
INGLÉS	27	HE CONOSCADO	CONOCER	DESV4
INGLÉS	27	HA ESTIDO	ESTAR	DESV4

**Tabla Nº 28: fragmento de Tabla dinámica 1 (inglés)**

**Fragmento de la tabla Nº 1 para el corpus holandés**

CORPUS	TEXTO	OCURRENCIA	VERBO	TIPO DESV_
HOLANDÉS	2	LOCALICÓ	LOCALIZAR	DESV3
HOLANDÉS	3	SOY	SER	EST
HOLANDÉS	3	SON	SER	EST
HOLANDÉS	3	HAY	HABER	EST
HOLANDÉS	3	ES	SER	EST
HOLANDÉS	3	PUEDE	PODER	EST
HOLANDÉS	3	SÉ	SABER	EST
HOLANDÉS	3	PUEDE	PODER	EST
HOLANDÉS	3	HAY	HABER	EST
HOLANDÉS	3	TIENE	TENER	EST
HOLANDÉS	3	SON	SER	EST
HOLANDÉS	3	HAY	HABER	EST
HOLANDÉS	3	PUEDEN	PODER	EST
HOLANDÉS	3	PUEDE	PODER	EST
HOLANDÉS	3	HAY	HABER	EST
HOLANDÉS	3	HAY	HABER	EST
HOLANDÉS	3	ES	SER	EST
HOLANDÉS	3	NACIÓ	NACER	EST
HOLANDÉS	4	MURIÓ	MORIR	EST



HOLANDÉS	4	ERA	SER	EST
HOLANDÉS	4	ERA	SER	EST
HOLANDÉS	4	TENÍA	TENER	EST
HOLANDÉS	4	TENÍA	TENER	EST
HOLANDÉS	4	TENÍA	TENER	EST
HOLANDÉS	4	ESTUVO	ESTAR	EST
HOLANDÉS	4	TENÍA	TENER	EST
HOLANDÉS	4	ERA	SER	EST
HOLANDÉS	4	HACÍA	HACER	EST
HOLANDÉS	4	ERA	SER	EST
HOLANDÉS	4	HAY	HABER	EST
HOLANDÉS	4	ES	SER	EST
HOLANDÉS	4	TENÍÓ	TENER	EST
HOLANDÉS	4	ENCONTRA	ENCONTRAR	DESV1
HOLANDÉS	4	CONTRIBURION	CONTRIBUIR	DESV1
HOLANDÉS	4	REQUIRE	REQUERIR	DESV3
HOLANDÉS	4	QUIERO	QUERER	DESV3
HOLANDÉS	5	ES	SER	EST
HOLANDÉS	5	HAY	HABER	EST
HOLANDÉS	5	FUE	SER	EST
HOLANDÉS	5	HAY	HABER	EST
HOLANDÉS	5	FUIMOS	IR	EST
HOLANDÉS	5	FUIMOS	IR	EST
HOLANDÉS	5	FUIMOS	IR	EST
HOLANDÉS	5	VOLVIMOS	VOLVER	EST
HOLANDÉS	5	TENIMOS	TENER	EST
HOLANDÉS	5	TENIMOS	TENER	DESV1
HOLANDÉS	5	TENIMOS	TENER	DESV1
HOLANDÉS	5	TENÍÓ	TENER	DESV1
HOLANDÉS	5	OFFRÓ	OFRECER	DESV1
HOLANDÉS	5	TENÍÓ	TENER	DESV3
HOLANDÉS	5	TENÍÓ	TENER	DESV4
HOLANDÉS	6	VOY	IR	DESV1
HOLANDÉS	7	PIENSO	PENSAR	EST
HOLANDÉS	7	TENEMOS	TENER	EST
HOLANDÉS	7	TENÍS	TENER	DESV4

Tabla Nº 29: fragmento de Tabla dinámica 1 (holandés)

Fragmento tabla Nº 1 corpus portugués

CORPUS	TEXTO	OCURRENCIA	VERBO	TIPO DESV-
PORTUGUÉS	1	TIENE	TENER	EST
PORTUGUÉS	1	PROPOSTO	PROPONER	DESV1
PORTUGUÉS	1	EXPONEMOS	EXPONER	EST
PORTUGUÉS	2	INCRUYO	INCLUIR	DESV3
PORTUGUÉS	3	ESTÁ	ESTAR	EST
PORTUGUÉS	3	ES	SER	EST
PORTUGUÉS	3	ESTÁ	ESTAR	EST
PORTUGUÉS	3	QUIERO	QUERER	EST
PORTUGUÉS	3	SALGA	SALIR	EST
PORTUGUÉS	4	ARREPINTINRON	ARREPENTIRSE	DESV3
PORTUGUÉS	4	ARREPENTIMOS	ARREPENTIRSE	EST
PORTUGUÉS	5	EXPONISTE	EXPONER	DESV1
PORTUGUÉS	5	PROPONAN	PROPONER	DESV4
PORTUGUÉS	6	PROPONÍ	PROPONER	DESV1
PORTUGUÉS	6	ESTOY	ESTAR	EST
PORTUGUÉS	6	ESTOY	ESTAR	EST
PORTUGUÉS	7	TIEM	TENER	DESV3
PORTUGUÉS	7	ENCUENTRAN	ENCONTRAR	EST
PORTUGUÉS	8	ES	SER	EST
PORTUGUÉS	8	ESTÁN	ESTAR	EST
PORTUGUÉS	8	ESTÁ	ESTAR	EST
PORTUGUÉS	8	HA	HABER	DESV3
PORTUGUÉS	8	SON	SER	EST
PORTUGUÉS	8	ESTÁN	ESTAR	EST
PORTUGUÉS	8	ESTÁ	ESTAR	EST
PORTUGUÉS	8	PARECEN	PARECER	EST
PORTUGUÉS	8	ESTÁN	ESTAR	EST
PORTUGUÉS	9	SON	SER	EST
PORTUGUÉS	9	DEMONSTRAN	DEMOSTRAR	DESV3
PORTUGUÉS	9	ES	SER	EST
PORTUGUÉS	10	ENCONTRASE	ENCONTRAR	DESV1
PORTUGUÉS	10	INCLUJENDO	INCLUIR	DESV3

Tabla Nº 30: fragmento de Tabla dinámica 1 (portugués)

A partir de esta base de datos por ocurrencia (siendo que cada fila es una palabra analizada) se confecciona una tabla dinámica llamada tabla 2. A su vez, esta base de datos, que nos brindará el total de ocurrencias por texto de cada corpus y el total de desviaciones por lengua y en general, será analizada estadísticamente.

Para ello se utiliza la opción *tabla dinámica* en Excel y se selecciona la tabla completa (todas las columnas y todas las filas) y se indica que los resultados se encuentren en una *nueva hoja de cálculo*. Luego se coloca la columna lengua como *filtro del informe*, la columna texto como *rótulo de fila* y la de tipo de desviación como *rótulo de columna*. De esta manera se hace visible la cantidad de ocurrencias por texto y el tipo de desviación.

**Tabla N ° 2. Base de datos obtenida (tabla dinámica)**

CORPUS	TEXTO	DESV1	DESV2	DESV3	DESV4	EST	TOTAL_V
ALEMÁN	1	1	1	4	2	12	20
ALEMÁN	2	0	0	0	0	2	2
ALEMÁN	3	2	0	1	0	10	13
ALEMÁN	4	0	0	0	1	16	17
ALEMÁN	5	0	0	0	0	1	1
ALEMÁN	6	1	0	0	1	3	5
ALEMÁN	7	0	0	0	0	2	2
ALEMÁN	8	4	0	0	0	0	4
ALEMÁN	9	1	0	0	1	0	2
ALEMÁN	10	2	0	0	0	0	2
ALEMÁN	11	0	0	0	1	5	6
ALEMÁN	12	0	0	2	0	2	4
ALEMÁN	13	0	0	0	1	20	21
ALEMÁN	14	0	0	0	0	9	9
ALEMÁN	15	0	0	0	1	25	26
ALEMÁN	16	0	0	0	0	7	7
ALEMÁN	17	0	0	0	0	7	7
ALEMÁN	18	0	0	0	0	13	13
ALEMÁN	19	0	0	0	0	8	8
ALEMÁN	20	0	0	0	0	11	11
ALEMÁN	21	0	0	0	0	17	17

FRANCÉS	1	0	0	0	0	5	5
FRANCÉS	2	0	0	0	0	1	1
FRANCÉS	3	0	0	0	0	5	5
FRANCÉS	4	0	0	0	0	6	6
FRANCÉS	5	0	0	0	0	3	3
FRANCÉS	6	0	0	0	0	4	4
FRANCÉS	7	0	0	0	0	3	3
FRANCÉS	8	0	0	0	0	4	4
FRANCÉS	9	4	0	3	1	3	11
FRANCÉS	10	0	0	0	0	2	2
FRANCÉS	11	1	0	0	0	0	1
FRANCÉS	12	0	0	0	0	4	4
FRANCÉS	13	0	0	0	0	2	2
FRANCÉS	14	0	0	0	0	3	3
FRANCÉS	15	0	0	1	1	5	7
FRANCÉS	16	0	0	0	0	2	2
FRANCÉS	17	0	0	0	0	3	3
FRANCÉS	18	0	0	0	0	7	7
FRANCÉS	19	0	0	0	0	4	4
FRANCÉS	20	0	0	0	1	4	5
FRANCÉS	21	0	0	0	0	5	5
FRANCÉS	22	0	0	0	0	1	1
FRANCÉS	23	0	0	0	0	5	5
FRANCÉS	24	0	0	0	0	9	9
FRANCÉS	25	0	0	0	0	8	8
FRANCÉS	26	0	0	0	2	0	2
FRANCÉS	27	0	0	1	0	2	3
FRANCÉS	28	0	0	0	0	4	4
FRANCÉS	29	1	1	0	0	8	10
FRANCÉS	30	0	1	0	1	9	11
FRANCÉS	31	0	0	0	0	13	13
FRANCÉS	32	0	0	0	0	18	18
FRANCÉS	33	1	0	0	0	1	2
FRANCÉS	34	0	0	0	0	10	10
FRANCÉS	35	0	0	0	0	1	1
FRANCÉS	36	0	0	0	0	5	5
FRANCÉS	37	0	0	0	0	7	7
FRANCÉS	38	0	0	1	0	3	4
FRANCÉS	39	0	0	0	0	4	4
FRANCÉS	40	0	0	0	0	10	10
FRANCÉS	41	0	0	0	0	7	7
FRANCÉS	42	0	0	0	0	11	11

FRANCÉS	43	0	0	0	1	6	7
FRANCÉS	44	0	0	0	1	5	6
FRANCÉS	45	0	0	0	0	4	4
FRANCÉS	46	0	0	0	0	6	6
FRANCÉS	47	0	0	0	0	3	3
FRANCÉS	48	0	0	0	0	6	6
HOLANDÉS	1	4	0	2	0	29	35
HOLANDÉS	2	2	0	2	0	5	9
HOLANDÉS	3	0	0	0	0	17	17
HOLANDÉS	4	2	0	2	0	14	18
HOLANDÉS	5	0	0	1	1	8	10
HOLANDÉS	6	1	0	0	1	0	2
HOLANDÉS	7	0	0	0	1	3	4
INGLÉS	1	0	0	0	0	4	4
INGLÉS	2	0	0	0	0	1	1
INGLÉS	3	0	1	0	0	4	5
INGLÉS	4	0	0	0	0	6	6
INGLÉS	5	1	0	0	0	3	4
INGLÉS	6	0	0	2	0	7	9
INGLÉS	7	0	0	0	0	1	1
INGLÉS	8	0	0	0	0	1	1
INGLÉS	9	0	0	0	0	2	2
INGLÉS	10	0	0	0	0	3	3
INGLÉS	11	0	0	0	0	1	1
INGLÉS	12	0	0	0	0	2	2
INGLÉS	13	0	0	0	0	4	4
INGLÉS	14	0	0	0	0	1	1
INGLÉS	15	0	0	0	0	5	5
INGLÉS	16	0	0	0	0	2	2
INGLÉS	17	0	0	0	0	1	1
INGLÉS	18	0	0	0	1	0	1
INGLÉS	19	0	0	0	0	1	1
INGLÉS	20	0	0	0	0	2	2
INGLÉS	21	0	0	0	0	2	2
INGLÉS	22	0	0	0	0	6	6
INGLÉS	23	2	0	1	0	12	15
INGLÉS	24	0	0	0	0	3	3
INGLÉS	25	1	1	0	0	3	5
INGLÉS	26	1	0	1	0	2	4
INGLÉS	27	2	0	1	2	14	19
INGLÉS	28	0	0	0	0	12	12
INGLÉS	29	0	0	0	0	7	7

INGLÉS	30	0	0	1	0	6	7
PORTUGUÉS	1	1	0	0	0	2	3
PORTUGUÉS	2	0	0	1	0	0	1
PORTUGUÉS	3	0	0	0	0	5	5
PORTUGUÉS	4	0	0	1	0	1	2
PORTUGUÉS	5	1	0	0	1	0	2
PORTUGUÉS	6	1	0	0	0	2	3
PORTUGUÉS	7	0	0	1	0	1	2
PORTUGUÉS	8	0	0	2	0	7	9
PORTUGUÉS	9	0	0	1	0	2	3
PORTUGUÉS	10	1	0	1	0	1	3
PORTUGUÉS	11	1	0	1	0	2	4
PORTUGUÉS	12	0	0	1	0	2	3
PORTUGUÉS	13	0	0	1	1	2	4
PORTUGUÉS	14	0	0	1	5	4	10
PORTUGUÉS	15	1	0	6	0	4	11
PORTUGUÉS	16	0	0	1	0	6	7
PORTUGUÉS	17	0	0	1	0	8	9
PORTUGUÉS	18	1	0	1	0	16	18
PORTUGUÉS	19	0	0	1	0	7	8
PORTUGUÉS	20	0	0	2	0	5	7
PORTUGUÉS	21	0	0	0	0	13	13
PORTUGUÉS	22	4	0	2	1	0	7
PORTUGUÉS	23	1	0	0	0	5	6
PORTUGUÉS	24	0	0	0	0	9	9
PORTUGUÉS	25	0	0	2	0	4	6
PORTUGUÉS	26	0	0	0	0	6	6
PORTUGUÉS	27	0	0	0	0	6	6

Tabla Nº 31: Tabla dinámica Nº2.

## 7.2. Técnica estadística no paramétrica en la comparación de corpus de distintas lenguas

La técnica estadística utilizada para realizar la comparación es la Prueba no paramétrica de Kruskal- Wallis.

El contraste de Kruskal-Wallis es la alternativa no paramétrica del método ANOVA, es decir, sirve para contrastar la hipótesis de que k muestras cuantitativas han sido obtenidas

de la misma población. Esta prueba es el método más adecuado para comparar poblaciones cuyas distribuciones no son normales. La única exigencia sobre los datos se refiere a la aleatoriedad en la extracción de las muestras, no haciendo referencia a ninguna de las otras condiciones necesarias para la validez de los resultados de una prueba paramétrica. Se basan en el uso de los rangos asignados a las observaciones.

Supongamos que se tienen  $k$  muestras de tamaños  $n_i$  respectivamente.

La hipótesis a contrastar puede enunciarse de la siguiente manera:

$$\begin{cases} H_0 : \text{Las } k \text{ muestras provienen de la misma población} \\ H_1 : \text{Alguna proviene de una población con mediana diferente a las demás} \end{cases}$$

El modo de realizar el contraste es el siguiente:

- Se ordenan las observaciones de menor a mayor y se asignan los rangos a cada una de ellas (1 para la menor, 2 para la siguiente,... y  $N$  para la observación mayor).
- Para cada una de las muestras, se calcula  $R_i$ , con  $i=1,2,...,k$  como la suma de los rangos de las observaciones de la muestra  $i$ . Si  $H_0$  es falsa, cabe esperar que esas cantidades sean muy diferentes.
- Se calcula el estadístico:

$$H = \frac{12}{N(N+1)} \sum_{i=1}^k \frac{R_i^2}{n_i} - 3(N+1)$$

- Regla de decisión:

Si  $H_0$  es falsa los valores de  $R_i$  serán muy diferentes y por lo tanto el valor de la estadística  $H$  será grande. Entonces, valores grandes de  $H$  estarán sustentando la hipótesis alternativa. Por lo tanto, el p-value será la probabilidad de obtener un valor de  $H$  igual o mayor que el observado, asumiendo que la hipótesis nula es cierta.

Cuando el valor de probabilidad asociada o p-value sea menor que el nivel de significación se rechazará la hipótesis nula.

Este procedimiento de análisis nos permite comparar la cantidad de desviaciones y de formas estándares en cuatro corpus distintos que en este caso poseen una cantidad similar de textos y de palabras, sin recurrir a supuestos respecto de la forma real de las distribuciones de probabilidad.<sup>17</sup> Se trabaja con el corpus francés que posee 48 textos y 6796 palabras, con el alemán de 21 textos y 2488 palabras, con el inglés con 30 textos y 1398 palabras y con el portugués de 1721 palabras y 27 textos. Debido a que el corpus holandés tiene una menor cantidad de datos, solo 7 textos, no se incluye en la comparación.

Se plantean dos hipótesis estadísticas:  $H_0$  o hipótesis nula que sostiene que las desviaciones son idénticas en todos los corpus, es decir que no hay diferencias significativas entre los mismos y la  $H_a$  o hipótesis alternativa que sostiene que al menos uno o dos de los corpus difiere del resto.

El modo de realizar el contraste es el siguiente: se ordenan los textos de menor a mayor, asignando a cada uno de ellos su rango (1 para el primero, 2 para el siguiente,  $N$  para el último). En el corpus francés  $N$  será 48, en el alemán 21 y en el inglés 30.

Para cada uno de los corpus, se calcula  $R_i$ ,  $i = 1, \dots, k$ , como la suma de los rangos de las observaciones que les corresponden. Si  $H_0$  es falsa, se espera que esas cantidades sean muy diferentes.

### Medidas resumen

CORPUS	Variablen	Media	D.E.	Mín	Máx	Mediana	
ALEMAN	PDESV1	21	13,83	30,94	0,00	100,00	0,00
ALEMAN	PDESV2	21	0,24	1,09	0,00	5,00	0,00
ALEMAN	PDESV3	21	3,70	11,56	0,00	50,00	0,00
ALEMAN	PDESV4	21	5,29	11,75	0,00	50,00	0,00

<sup>17</sup> Mendenhall, W. Wackerly, D. Scheaffer, R. (2002) *Estadística matemática con aplicaciones*. Grupo Editorial Iberoamérica, México.



ALEMAN	PEST	21	76,94	35,65	0,00	100,00	96,15
FRANCÉS	PDESV1	48	4,61	16,96	0,00	100,00	0,00
FRANCÉS	PDESV2	48	0,40	1,93	0,00	10,00	0,00
FRANCÉS	PDESV3	48	1,56	6,43	0,00	33,33	0,00
FRANCÉS	PDESV4	48	3,82	14,98	0,00	100,00	0,00
FRANCÉS	PEST	48	89,61	23,87	0,00	100,00	100,00
HOLANDÉS	PDESV1	7	24,49	34,79	0,00	100,00	11,43
HOLANDÉS	PDESV2	7	0,00	0,00	0,00	0,00	0,00
HOLANDÉS	PDESV3	7	6,53	8,12	0,00	22,22	5,71
HOLANDÉS	PDESV4	7	5,71	12,43	0,00	33,33	0,00
HOLANDÉS	PEST	7	63,27	31,70	0,00	100,00	66,67
INGLÉS	PDESV1	30	2,46	6,84	0,00	25,00	0,00
INGLÉS	PDESV2	30	2,00	8,05	0,00	40,00	0,00
INGLÉS	PDESV3	30	2,45	6,47	0,00	25,00	0,00
INGLÉS	PDESV4	30	3,68	18,29	0,00	100,00	0,00
INGLÉS	PEST	30	89,41	21,61	0,00	100,00	100,00
PORTUGUÉS	PDESV1	27	9,73	17,04	0,00	57,14	0,00
PORTUGUÉS	PDESV2	27	0,31	1,60	0,00	8,33	0,00
PORTUGUÉS	PDESV3	27	16,63	23,73	0,00	100,00	10,00
PORTUGUÉS	PDESV4	27	5,16	14,02	0,00	50,00	0,00
PORTUGUÉS	PEST	27	68,17	31,48	0,00	100,00	75,00

### 7.2.1. Resultados

#### 1. DESVIACIÓN 1 (regularizar una forma irregular)

Variable CORPUS	N	Medias	D.E.	Medianas	H	p
PDESV1 ALEMAN	21	13,83	30,94	0,00	3,58	<b>0,0542</b>
PDESV1 FRANCES	48	4,61	16,96	0,00		
PDESV1 INGLES	30	2,46	6,84	0,00		
PDESV1 PORTUGUÉS	27	9,73	17,04	0,00		

**NO SE OBSERVAN DIFERENCIAS SIGNIFICATIVAS RESPECTO DEL PORCENTAJE DE DESV1 ENTRE LOS CORPUS**

## 2. DESVIACIÓN 2 (Hacer irregular una forma que es regular)

Variable	CORPUS	N	Medias	D.E.	Medianas	H	p
PDESV2	ALEMAN	21	0,24	1,09	0,00	0,06	<b>0,9382</b>
PDESV2	FRANCES	48	0,40	1,93	0,00		
PDESV2	INGLES	30	2,00	8,05	0,00		
PDESV2	PORTUGUÉS	27	0,31	1,60	0,00		

Página | 82

**NO SE OBSERVAN DIFERENCIAS SIGNIFICATIVAS RESPECTO DEL PORCENTAJE DE DESV2 ENTRE LOS CORPUS**

## 3. DESVIACIÓN 3 (Formas verbales que presentan una irregularidad distinta a la estándar)

Variable	CORPUS	N	Medias	D.E.	Medianas	H	p
PDESV3	ALEMAN	21	3,70	11,56	0,00	13,71	<b>&lt;0,0001</b>
PDESV3	FRANCES	48	1,56	6,43	0,00		
PDESV3	INGLES	30	2,45	6,47	0,00		
PDESV3	PORTUGUÉS	27	16,63	23,73	10,00		

Trat.	Ranks	
FRANCES	54,58	A
ALEMAN	59,50	A
INGLES	60,22	A
PORTUGUÉS	86,11	B

*Medias con una letra común no son significativamente diferentes ( $p > 0,05$ )*

**SE OBSERVAN DIFERENCIAS SIGNIFICATIVAS RESPECTO DEL PORCENTAJE DE DESV3 ENTRE LOS CORPUS.**

**LAS COMPARACIONES MÚLTIPLES SIGUIENTES EVIDENCIAN QUE EL CORPUS DE PORTUGUÉS PRESENTA VALORES SIGNIFICATIVAMENTE SUPERIORES A LOS RESTANTES. ENTRE LOS OTROS TRES NO HAY DIFERENCIAS.**

## 4. DESVIACIÓN 4 (Asignación incorrecta de desinencia)

Variable	CORPUS	N	Medias	D.E.	Medianas	H	p
PDESV4	ALEMÁN	21	5,29	11,75	0,00	2,33	<b>0,1235</b>
PDESV4	FRANCÉS	48	3,82	14,98	0,00		
PDESV4	INGLÉS	30	3,68	18,29	0,00		
PDESV4	PORTUGUÉS	27	5,16	14,02	0,00		

**NO SE OBSERVAN DIFERENCIAS SIGNIFICATIVAS RESPECTO DEL PORCENTAJE DE DESV4 ENTRE LOS CORPUS**

## 5. Formas coincidentes con la lengua estándar

Variable	CORPUS	N	Medias	D.E.	Medianas	H	p
PEST	ALEMAN	21	76,94	35,65	96,15	15,00	0,0003
PEST	FRANCES	48	89,61	23,87	100,00		
PEST	INGLES	30	89,41	21,61	100,00		
PEST	PORTUGUÉS	27	68,17	31,48	75,00		

Trat.	Ranks
PORTUGUÉS	42,11 A
ALEMAN	56,86 A B
INGLES	71,08 B
FRANCES	73,70 B

*Medias con una letra común no son significativamente diferentes ( $p > 0,05$ )*

**SE OBSERVAN DIFERENCIAS SIGNIFICATIVAS RESPECTO DEL PORCENTAJE DE EST ENTRE LOS CORPUS. LAS COMPARACIONES MÚLTIPLES SIGUIENTES EVIDENCIAN QUE EL CORPUS DE PORTUGUÉS PRESENTA VALORES SIGNIFICATIVAMENTE INFERIORES AL CORPUS DE INGLÉS Y FRANCÉS. ENTRE LOS OTROS TRES NO HAY DIFERENCIAS.**

La regla para decidir si se ha de rechazar o no la hipótesis nula es la siguiente: se rechaza  $H_0$  si  $p$  es menor a 0,05. Como los resultados obtenidos de  $p$  son mayores a 0,05 teniendo: **0,0542** para la DESV1; **0,9382** para la DESV2 y **0,1235** para la DESV4,  $H_0$  no se rechaza y por lo tanto, los cuatro corpus comparados no difieren respecto del porcentaje de desviaciones del tipo 1, 2 y 4.

Con respecto a la DESV3, hallamos que el corpus de portugués nos arroja un resultado menor:  **$p < 0,0001$** , y por lo tanto presenta diferencias significativas dadas por una cantidad mayor de desviaciones. Los otros corpus no presentan diferencias entre ellos.

En cuanto a las formas de la lengua estándar (EST) también es el corpus portugués el que se distingue del resto por presentar diferencias evidenciando una cantidad inferior de formas coincidentes con el español.

De la lectura de estos datos interpreto que el corpus portugués presenta una mayor cantidad de formas desviadas respecto al resto de los corpus. A la vez, este hecho se evidencia en la menor proporción de formas coincidentes con la lengua meta que exhibe.

La explicación de esta diferencia podría deberse a que los sujetos pertenecientes al corpus portugués se encuentran en un nivel básico de adquisición del español pero inferior al resto de los aprendientes debido a que no se hallan inmersos en una comunidad lingüística de habla española siendo que sus estancias en Argentina son cortas. Esto es lo que los distingue de los otros sujetos pertenecientes a los corpus inglés, francés y alemán que sí se encuentran viviendo en el país y por lo tanto están expuestos a una mayor cantidad de datos.

### 7.3. Frecuencia de los verbos

Creamos la tabla dinámica a la que distinguiremos con el número 3 para poder obtener resultados acerca de cuáles fueron los verbos más utilizados en los cuatro corpus y cuáles de ellos presentan mayor cantidad de desviaciones. Para esto insertamos una nueva tabla dinámica, a la cual se le indica que como *filtro de informe* muestre el corpus, como *rótulo de fila* coloque los verbos irregulares y por último, como *rótulo de columna* ubique al tipo de desviación. De esta manera, y llevando el valor ocurrencia hasta el centro de la tabla, le ordenamos a la herramienta que sume la cantidad de desviaciones halladas en un total general por verbo.

Tabla Nº 3 de frecuencia de verbos

Etiquetas de fila	DESV1	DESV2	DESV3	DESV4	EST	Total general
AGREGAR			1			1
AMENAZAR					1	1

ANDAR				2	2
ARREPENTIRSE		1		1	2
BUSCAR		1			1
CALENTAR	1				1
CARECER				1	1
COLOCAR	1	1			2
COMENZAR	1				1
CONducir				1	1
CONocer		2	1	9	12
CONSTRUIR				1	1
CONTAR	1				1
CONTRIBUIR	1				1
COSTAR		1			1
DAR		1		6	7
DECIR				15	15
DEMOSTRAR		3			3
DERRETIR				1	1
DESCUBRIR				1	1
DESEMBARCAR		1			1
DETENER	1	1			2
DIRIGIR		1			1
DOLER				1	1
DORMIR				3	3
ELEGIR				1	1
EMBARCAR		1			1
EMPEZAR	2	1	1	6	9
ENCONTRAR	5	1	3	10	19
ESTAR		3	9	80	92
EXPONER	1			1	2
FREIR	1			1	2
HABER	2	1	1	38	42
HACER	4	2		19	25
INCLUIR		2			2
IR	1	1	1	52	55
JUGAR	1				1
LLEGAR	1	1		2	4
LOCALIZAR		1			1
MOLER	1				1
MORIR				1	1
MOSTRAR		1		1	2
MOVER				1	1
NACER				3	3

OFRECER	1			1	2	
PAGAR			1	1	1	3
PARECER			1		10	11
PEDIR	2				2	4
PENSAR	2		1		2	5
PERCIBIR			1			1
PICAR	1		1			2
PODER	1	1		1	38	41
PONER			1		6	7
PREFERIR	1			1	1	3
PROPONER	3			1		4
QUERER	1	2	1	1	65	70
RECOMENDAR					3	3
RECORDAR					2	2
REENCONTRARSE					1	1
RENOVAR	1					1
REQUERIR			1			1
ROGAR		1		1		2
ROMPER					1	1
SABER					10	10
SACAR	2			1	1	4
SALIR					3	3
SENTARSE					1	1
SENTIR				1	12	13
SER					196	196
SOLER					1	1
SONREÍR					1	1
SUGERIR	1				1	2
SUPONER					2	2
TENER	7		5	4	94	110
TRAER	1		2		2	5
VENIR	1				6	7
VER				1	9	10
VOLVER	1	1		1	2	5
Total general	52	7	47	28	733	867

**Tabla Nº 32: Frecuencia de verbos irregulares**

Los verbos irregulares más frecuentes en la interlengua son los siguientes:

VERBO	FREC
<b>SER</b>	196
<b>TENER</b>	110
<b>ESTAR</b>	92
<b>QUERER</b>	70
<b>IR</b>	55
<b>HABER</b>	42

**Tabla Nº 33: Verbos irregulares más frecuentes**

Los verbos que presentan mayor cantidad de desviaciones en la interlengua son los siguientes:

VERBO	DESV
<b>TENER</b>	16
<b>ESTAR</b>	12
<b>ENCONTRAR</b>	9

**Tabla Nº 34: verbos con mayor cantidad de desviaciones**

#### Corpus alemán:

CORPUS	ALEMAN					
Cuenta de OCURRENCIA	Etiquetas de columna					Total
Etiquetas de fila	DESV1	DESV2	DESV3	DESV4	EST	general
ANDAR					1	1
BUSCAR			1			1
CARECER					1	1
CONOCER					5	5
DAR					4	4
DECIR					13	13
DOLER				1		1
DORMIR					1	1
ELEGIR					1	1

EMPEZAR				5	5
ENCONTRAR		1		2	3
ESTAR			1	12	13
HABER			1	7	8
HACER	1	1		8	10
IR				15	15
JUGAR	1				1
PAGAR		1	1	1	3
PARECER				1	1
PEDIR	2			1	3
PODER				6	6
PONER		1		1	2
PREFERIR	1		1		2
PROPONER	1				1
QUERER			1	5	6
RECOMENDAR				2	2
RECORDAR				2	2
ROGAR		1	1		2
ROMPER				1	1
SABER				8	8
SACAR	1			1	2
SALIR				1	1
SENTARSE				1	1
SENTIR			1	5	6
SER				27	27
SOLER				1	1
SUGERIR	1			1	2
TENER				23	23
TRAER	1	2		2	5
VENIR	1			2	3
VER				2	2
VOLVER	1			1	2
<b>Total general</b>	<b>11</b>	<b>1</b>	<b>7</b>	<b>8 170</b>	<b>197</b>

Tabla Nº 35: frecuencia de verbos irregulares (alemán)



Los verbos más frecuentes en el corpus alemán son:

VERBO	FREC
SER	27
TENER	23
IR	15
DECIR	13
ESTAR	13

Tabla Nº 36: verbos con mayor frecuencia (alemán)

Con respecto al verbo que presenta mayor cantidad de desviaciones hallamos que es **traer**.

#### Corpus francés:

CORPUS	FRANCÉS					
Cuenta de OCURRENCIA		Etiquetas de columna				
Etiquetas de fila	DESV1	DESV2	DESV3	DESV4	EST	Total general
AGREGAR			1			1
AMENAZAR					1	1
CALENTAR	1					1
COLOCAR	1		1			2
COMENZAR	1	1				1
CONducir					1	1
CONocer			1		1	2
CONSTRUIR					1	1
DAR					1	1
DECIR					2	2
DERRETIR					1	1
DESCUBRIR					1	1
DIRIGIR			1			1
EMPEZAR	2				1	3
ENCONTRAR					5	5
ESTAR					1 25	26

FREIR	1		1		2
HABER				10	10
HACER				7	7
IR				20	20
MOSTRAR				1	1
NACER				2	2
PARECER				4	4
PEDIR				1	1
PENSAR				1	1
PICAR	1	1			2
PODER			1	16	17
PONER				4	4
QUERER				30	30
RECOMENDAR				1	1
RENOVAR	1				1
SACAR	1		1		2
SER				68	68
SUPONER				1	1
TENER			2	28	30
VENIR				2	2
VER			1	5	6
VOLVER		1	1		2
<b>Total general</b>	<b>9</b>	<b>2</b>	<b>5</b>	<b>8 243</b>	<b>268</b>

Tabla N° 37: frecuencia de verbos (francés)

Los verbos más utilizados en el corpus francés son:

VERBO	FREC
<b>SER</b>	68
<b>QUERER</b>	30
<b>TENER</b>	30
<b>ESTAR</b>	26

Tabla N° 38: verbos con mayor frecuencia (francés)

No hay verbos que se destaquen por presentar una mayor cantidad de desviaciones.

**Corpus inglés:**

CORPUS	INGLÉS					
Cuenta de OCURRENCIA	Etiquetas de columna					
Etiquetas de fila	DESV1	DESV2	DESV3	DESV4	EST	Total general
CONOCER			1	1	1	3
DAR			1		1	2
DORMIR					2	2
ESTAR				1	6	7
HABER	2					2
HACER	2		1		2	5
IR			1	1	11	13
PENSAR	1		1			2
PODER		1			11	12
PONER					1	1
PREFERIR					1	1
QUERER	1	2			28	31
SENTIR					1	1
SER					40	40
TENER			1		12	13
VENIR					1	1
Total general	6	3	6	3	118	136

**Tabla N° 39: frecuencia de verbos (inglés)**

Los verbos más frecuentes del corpus inglés se muestran en el siguiente cuadro:

VERBO	FREC
<b>TENER</b>	40
<b>QUERER</b>	31
<b>IR</b>	13
<b>PODER</b>	12

**Tabla N° 40: verbos con mayor frecuencia (inglés)**

**Hacer** y **querer** son los verbos que presentan mayor cantidad de desviaciones.

**Corpus portugués:**

CORPUS		PORTUGUÉS				
Cuenta de OCURRENCIA		Etiquetas de columna				
Etiquetas de fila	DESV1	DESV2	DESV3	DESV4	EST	Total general
ANDAR					1	1
ARREPENTIRSE			1		1	2
CONOCER					2	2
COSTAR			1			1
DEMOSTRAR			2			2
DESEMBARCAR			1			1
EMBARCAR			1			1
EMPEZAR			1			1
ENCONTRAR	3	1	2		3	9
ESTAR			3	6	34	43
EXPONER	1				1	2
HABER			1		9	10
INCLUIR			2			2
IR					3	3
LLEGAR	1		1		2	4
OFRECER					1	1
PARECER			1		5	6
PENSAR	1					1
PERCIBIR			1			1
PROPONER	2			1		3
QUERER					2	2
REENCONTRARSE					1	1
SABER					1	1
SALIR					2	2
SENTIR					6	6
SER					33	33
SONREÍR					1	1
SUPONER					1	1

TENER	4		3	14	21
VENIR				1	1
VER				2	2
<b>Total general</b>	<b>12</b>	<b>1</b>	<b>21</b>	<b>7 126</b>	<b>167</b>

**Tabla Nº 41: frecuencia de verbos (portugués)**

Los verbos más frecuentes en el corpus portugués son:

VERBO	FREC
<b>ESTAR</b>	43
<b>SER</b>	33
<b>TENER</b>	21

**Tabla Nº 42: verbos con mayor frecuencia (portugués)**

**Estar** y **tener** son los que presentan mayor cantidad de desviaciones.

**Corpus holandés:**

CORPUS	HOLANDÉS				
Cuenta de OCURRENCIA		Etiquetas de columna			
Etiquetas de fila	DESV1	DESV3	DESV4	EST	Total general
CONTAR	1				1
CONTRIBUIR	1				1
DEMOSTRAR		1			1
DETENER	1	1			2
ENCONTRAR	2				2
ESTAR				3	3
HABER				12	12
HACER	1			2	3
IR	1			3	4
LOCALIZAR		1			1

MOLER	1				1
MORIR				1	1
MOSTRAR		1			1
MOVER				1	1
NACER				1	1
OFRECER	1				1
PENSAR				1	1
PODER	1			5	6
QUERER		1			1
REQUERIR		1			1
SABER				1	1
SER				28	28
TENER	3	1	2	17	23
VOLVER				1	1
Total general	13	7	2	76	98

Tabla Nº 43: frecuencia de verbos (holandés)

Los verbos más utilizados en el corpus holandés son los siguientes:

VERBO	FREC
SER	28
TENER	23
HABER	12

Tabla Nº 44: verbos con mayor frecuencia (holandés)

Por su parte, **estar** y **tener** son los que presentan mayor cantidad de desviaciones.

## 7.4 Resultados

### Porcentaje de las formas verbales irregulares utilizadas en cada corpus

Página | 95

	CORPUS FRANCÉS	CORPUS ALEMÁN	CORPUS INGLÉS	CORPUS PORTUGUÉS
<b>TOTAL PALABRAS</b>	6796	2488	1398	1721
<b>FORMAS VERBALES IRREGULARES</b>	268	197	136	167
<b>PORCENTAJES</b>	<b>3,88%</b>	<b>7,91%</b>	<b>9,72%</b>	<b>9,70%</b>

Tabla Nº 45: Porcentaje de formas verbales irregulares

### Porcentaje según el tipo de desviación

	TOTAL OCURRENCIAS	DESV1	DESV2	DESV3	DESV4	EST
	<b>867</b>	<b>52</b>	<b>7</b>	<b>47</b>	<b>28</b>	<b>733</b>
<b>PORCENTAJES</b>	<b>100%</b>	<b>5,99%</b>	<b>0,80%</b>	<b>5,31%</b>	<b>3,22%</b>	<b>84,54%</b>

Tabla Nº 46: Porcentaje de formas verbales según tipo de desviación

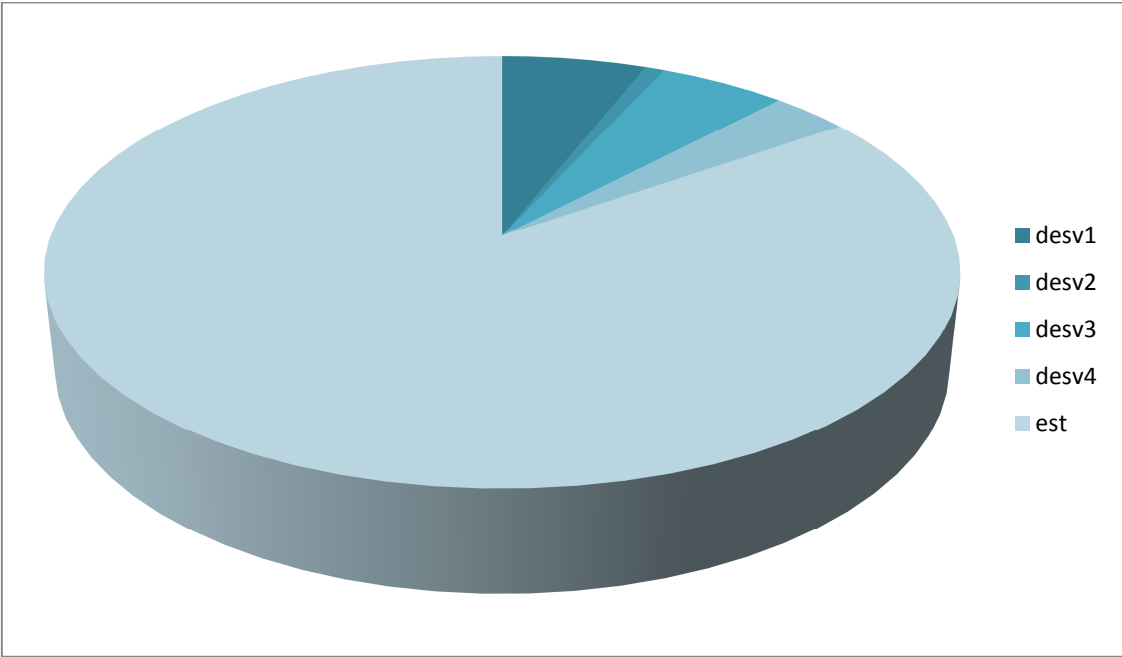


Gráfico Nº 1: Porcentajes desviaciones en todas las muestras.

Desviaciones por corpus

Corpus francés						Total
TIPO DESV_	DESV1	DESV2	DESV3	DESV4	EST	
	9	2	6	8	243	268
PORCENTAJE	3,37%	0,74%	2,24%	2,99%	90,67%	100%

TOTAL			
FORMAS DESVIADAS		9%	ESTÁNDAR
			91%

Tabla Nº 47: Porcentaje de formas desviadas corpus francés



En el gráfico de torta que sigue se muestran los porcentajes de desviaciones del corpus francés:

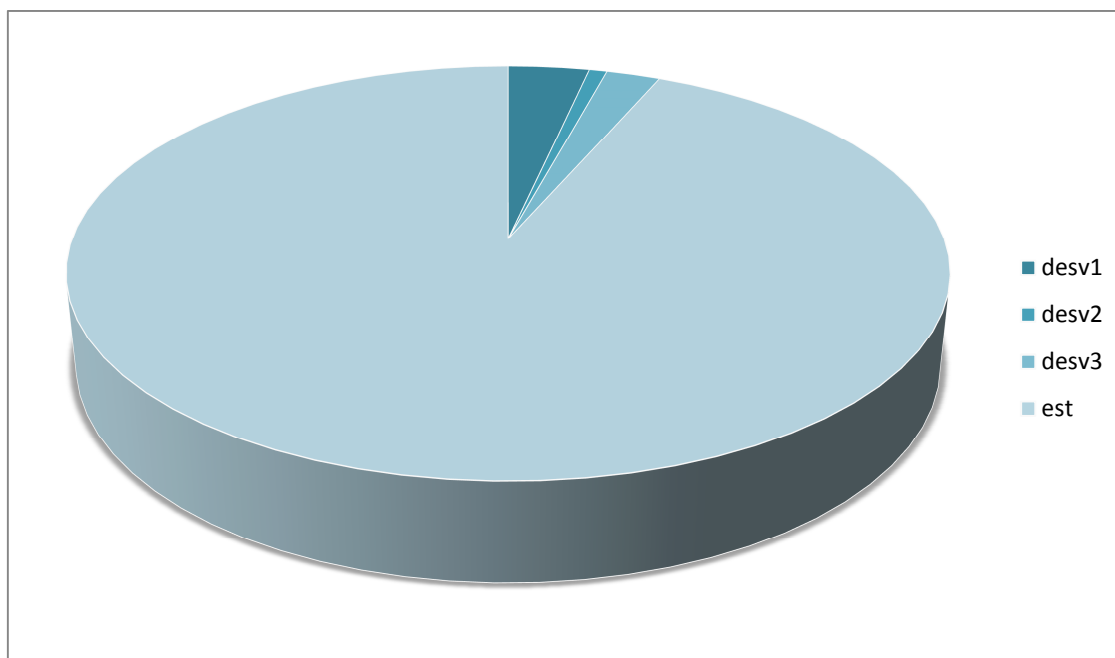


Gráfico Nº 2: desviaciones corpus francés

Corpus alemán						Total
TIPO DESV_	DESV1	DESV2	DESV3	DESV4	EST	
	11	1	7	8	170	197
PORCENTAJE	5,58%	0,50%	3,55%	4,06%	86,29%	100%
TOTAL						
FORMAS DESVIADAS			ESTÁNDAR			
14%			86%			

Tabla Nº 48: Porcentaje de formas desviadas corpus alemán

En el siguiente gráfico se muestran los porcentajes de las desviaciones:

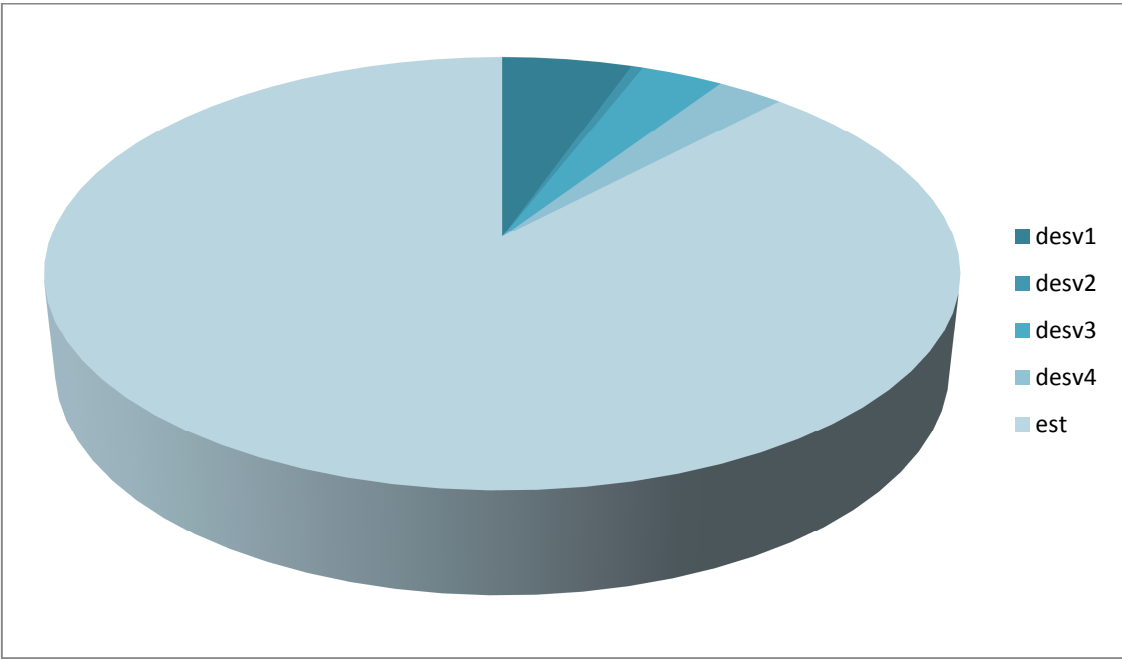


Gráfico Nº 3: Desviaciones corpus alemán

Corpus inglés						Total
TIPO DESV_	DESV1	DESV2	DESV3	DESV4	EST	
	7	2	6	3	118	136
PORCENTAJE	5,14%	1,47%	4,41%	2,20%	86,76%	100%

TOTAL					
FORMAS DESVIADAS		13%	ESTÁNDAR		87%

Tabla Nº 49: Porcentaje de formas desviadas corpus inglés

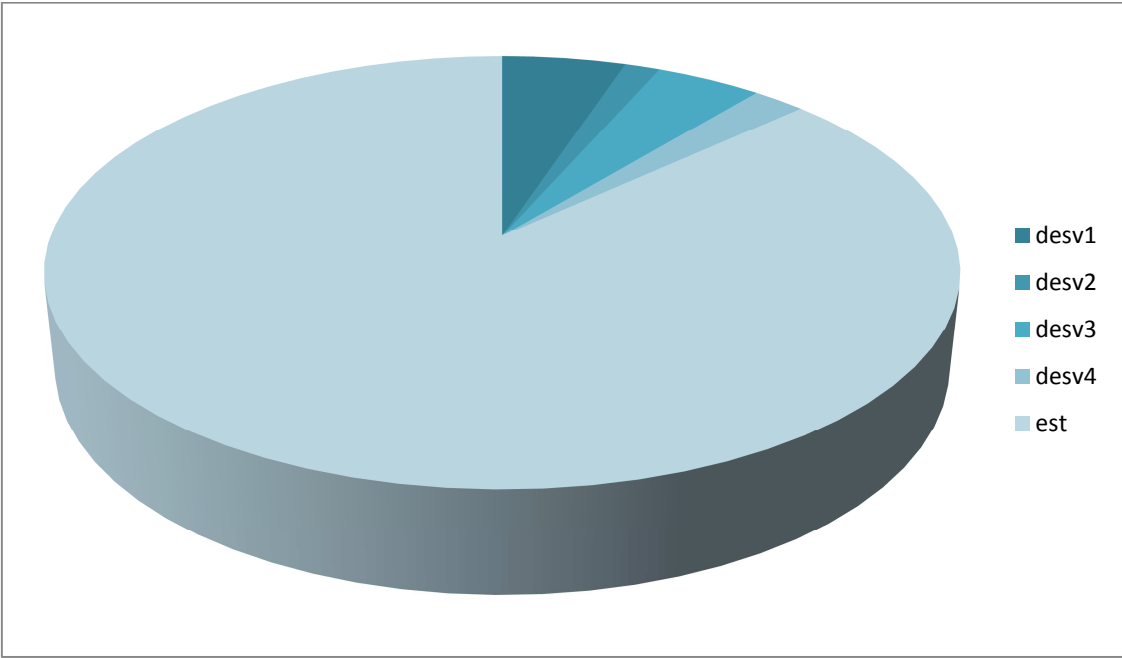


Gráfico Nº 4: Desviaciones corpus inglés.

Corpus portugués						Total
TIPO DESV_	DESV1	DESV2	DESV3	DESV4	EST	
	12	1	21	7	126	167
PORCENTAJE	7,14%	0,59%	12,5%	4,19%	75%	100%

TOTAL					
FORMAS DESVIADAS		25%	ESTÁNDAR		
			75%		

Tabla Nº 50: Porcentaje de formas desviadas corpus portugués

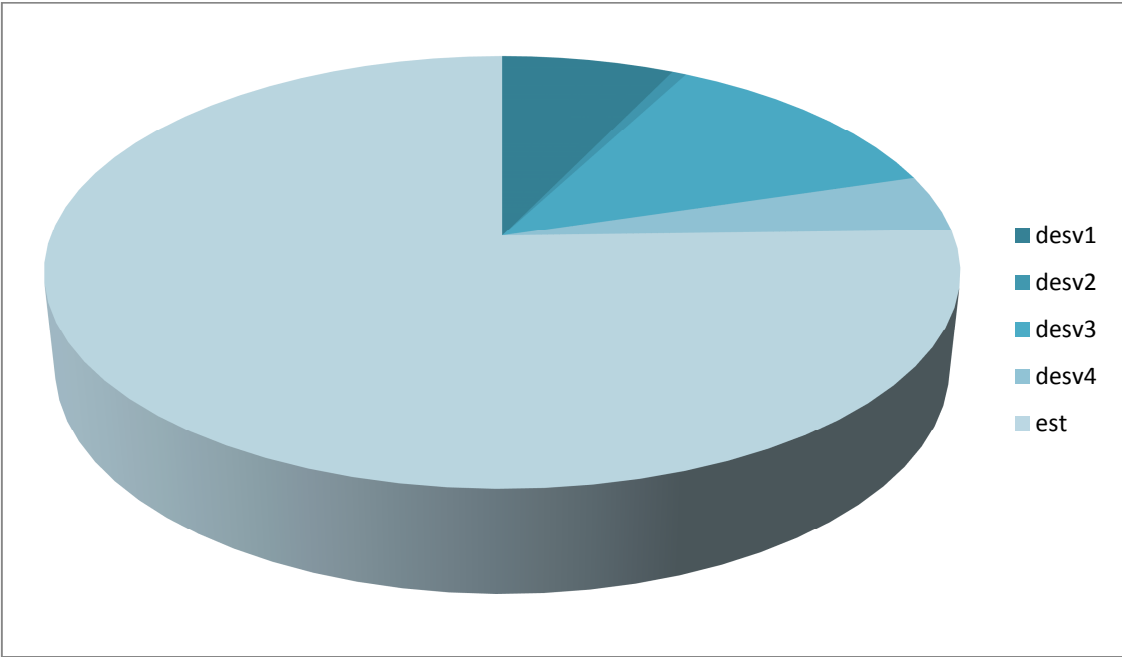
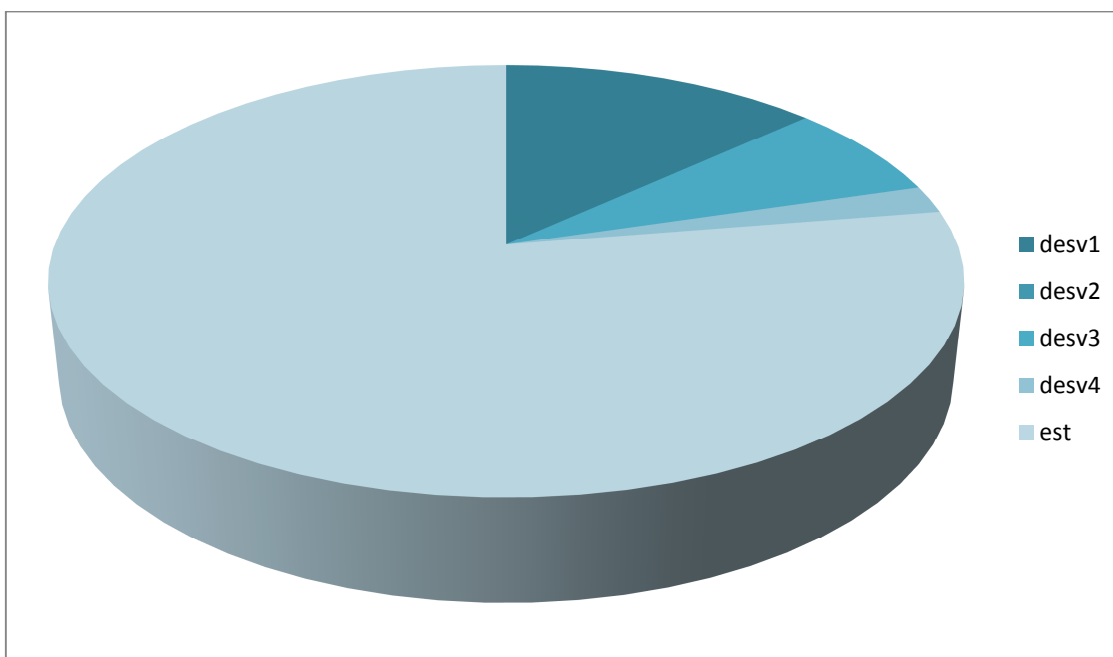


Gráfico Nº 5: Desviaciones corpus portugués.

Corpus holandés						Total
TIPO DESV_	DESV1	DESV2	DESV3	DESV4	EST	
	13	----	7	2	76	98
PORCENTAJE	13,26%	0%	7,14%	2,04%	77,55%	100%

Tabla Nº 51: Porcentaje de formas desviadas corpus holandés



**Gráfico N° 6. Desviaciones corpus holandés.**

## 7.5. Conclusiones

Luego se utilizó Excel para confeccionar la tabla 1 que reúne los siguientes datos: cada una de las ocurrencias por lengua, el verbo del que provienen y el tipo de desviación que presentan en caso de que no sean coincidentes con la lengua estándar.

A partir de ésta se creó una tabla dinámica en Excel, designada tabla 2, que funciona como base de datos y a partir de la cual se obtuvo el total de ocurrencias y el total de desviaciones tanto en cada lengua en particular como en todos los corpus en general.

Para el análisis estadístico se utilizó la prueba no paramétrica de Kruskal- Wallis, la cual permitió comparar los tres corpus sin recurrir a supuestos y dio como resultado que los mismos no difieren respecto del porcentaje de desviaciones del tipo 1, 3 y 4.

Con respecto al porcentaje de desviaciones del tipo 3 y estándar, al evidenciar la prueba, diferencias significativas, se procedió a realizar un test de comparaciones múltiples que arrojó como resultado que el corpus portugués es el que presenta una mayor cantidad de desviaciones del tipo 3 y una menor cantidad de formas estándar con respecto al resto de los corpus. Es importante aclarar, que los otros corpus: alemán, francés e inglés no presentan diferencias significativas entre sí en lo que respecta a dichos tipos de ocurrencias.

Por último se realizó otra tabla dinámica identificada con el número 3, con el fin de determinar cuáles de los verbos irregulares habían sido utilizados en mayor medida y cuáles de ellos presentaban una cantidad de desviaciones superior al resto.

Los resultados fueron los siguientes: el verbo ser es el más utilizado, seguido por el verbo tener. Podemos pensar que se debe al hecho de que se emplean frecuentemente en las narraciones y descripciones, teniendo en cuenta los tipos de textos que forman parte del corpus de interlengua.

Con respecto a los verbos irregulares que muestran mayor cantidad de desviaciones: en primer lugar hallamos el verbo tener (que sufre cambios consonánticos y vocálicos en el presente: tengo-tiene, es de pretérito fuerte: tuve-tuvieron y además varía en el futuro: tendré). En segundo lugar, el verbo estar (presenta la particularidad en presente de flexionarse como estoy y es de pretérito fuerte: estuve-estuvieron). Recordemos que todos ellos son verbos que presentan irregularidades especiales según la clasificación propuesta.

Finalmente, se pudo comprobar que el corpus portugués es el que posee un mayor porcentaje de formas verbales desviadas de los cuatro corpus comparados: 25%. El corpus alemán 14%, muy similar al porcentaje del corpus inglés: 13%, mientras que el francés es el que exhibe una cifra menor: 9%.

En esta comparación no se toman en cuenta los datos del corpus holandés debido a que posee una cantidad inferior de textos.

En cuanto al tipo de desviación más frecuente si reunimos todos los corpus, tenemos en primer lugar la desviación que refiere a hacer regular las formas irregulares, le sigue la que implica una irregularidad distinta a la estándar, y luego la que agrupa a las formas que exhiben una asignación incorrecta de desinencia.

En último lugar hallamos la desviación que alude al hecho de hacer irregular una forma que es regular.

Los programas informáticos que se utilizarán para poder realizar el reconocimiento automático de las formas verbales irregulares, considerando tanto las formas que coinciden con el español estándar como las desviadas, son dos: Smorph y Nooj<sup>18</sup>.

¿Por qué dos herramientas? La respuesta refiere a que las primeras pruebas, en el comienzo de la investigación, fueron realizadas mediante el programa Smorph<sup>19</sup>, un analizador y generador morfosintáctico, que realiza la tokenización y el análisis morfológico en una sola etapa, obteniendo como resultado las formas correspondientes a un lema con los valores adecuados. Al avanzar con este trabajo comencé a estudiar una nueva herramienta, de libre acceso y disponible on line, me refiero al sistema Nooj, que es una herramienta para el tratamiento de lenguas naturales desarrollada por Max Silberstein en 2002.

Por lo tanto, en este capítulo voy a comparar el funcionamiento de una y otra herramienta para mostrar los alcances que tienen en cuanto al objetivo propuesto, que es el de analizar automáticamente la interlengua de aprendientes de español, tratando de hallar la manera más económica de efectuar el reconocimiento de las formas verbales que presentan desviaciones respecto de la lengua meta.

Una de las ventajas de Nooj es que puede procesar textos y corpus de cientos de archivos de texto, además representa una herramienta multilingüe. Entre las 20 lenguas referidas encontramos: inglés, francés, portugués, alemán, italiano y ruso.

Para el módulo español (Argentina), mi equipo de trabajo es quien está ocupándose de cargar los datos lingüísticos correspondientes a gramáticas y diccionarios, creando los modelos y rasgos correspondientes a las diferentes clases gramaticales<sup>20</sup>, siguiendo las indicaciones de Max Silberstein.

<sup>18</sup> Programa de libre acceso, creado por Max Silberstein en 2002. disponible en la página:

[http://www.nooj-association.org/index.php?option=com\\_nooj&controller=application&task=display\\_application\\_fo&Itemid=570](http://www.nooj-association.org/index.php?option=com_nooj&controller=application&task=display_application_fo&Itemid=570)

<sup>19</sup> Creado en el GRILL por Salah Ait-Mokhtar bajo la dirección de Gabriel Bés.

<sup>20</sup> Ver publicaciones en la Bibliografía:

Bonino, R. (2011) "Una propuesta para la implantación de la morfología verbal del español" en Revista *INFOSUR* Nº 5, págs. 79-86.  
Solana, Zulema; Beltrán, Celina; Tramallino Carolina (2013) "Análisis automático morfológico con las herramientas SMORPH Y NOOJ" (En *Revista de epistemología y ciencias humanas*. Grupo IANUS, Docentes de la Universidad Nacional de Rosario y la Universidad del Litoral. ISSN 1852-625X. Nº 5, abril de 2013.

Tramallino, C. (2013) "Análisis morfológico con herramientas informáticas. Reconocimiento de nombres en textos de español con el sistema Nooj" en *Revista Lingüística y Literatura* ISSN 0120-5587. Universidad de Antioquía, Medellín, Colombia. Año 34, Nro 64 (julio-diciembre 2013) ISSN 01205587. Disponible en <http://aprendeonline.udea.edu.co/revistas/index.php/lyl>



Para ello se procedió a adaptar y actualizar la información lingüística declarada en los archivos de texto (.txt) del software Smorph<sup>21</sup>. El equipo Infosur declaró las siguientes fuentes lingüísticas: 4000 verbos, 28000 nombres, 8000 adjetivos, preposiciones, conjunciones y adverbios. La declaración comprende las entradas (o diccionario), los modelos, los rasgos y las terminaciones.

A partir de esas fuentes consignadas se procedió a trasladar la información contenida en los archivos de Smorph al sistema Nooj, previa adaptación de los algoritmos requeridos. Tal propósito es viable ya que las gramáticas y diccionarios de Nooj son fáciles de crear y de adecuarse. Ambos programas incluyen herramientas que sirven para chequear, limpiar, modificar, conservar y compartir los archivos.

El aporte personal reside en poder realizar una adaptación de la formalización lingüística empleada en estos programas con la finalidad de poder ejecutar el análisis, la búsqueda de información y la comprobación de hipótesis en textos de español producidos por aprendientes de segunda lengua.

### 8.1. Smorph

Smorph fue desarrollado en el Gril (Groupe de Recherche dans les Industries de la Langue), bajo la dirección de Gabriel G. Bès, este software no solo permite el análisis morfológico y el etiquetado de los textos, sino que también tiene en cuenta el análisis sintáctico. Además, en el caso de palabras desconocidas o no reseñadas en el diccionario, consigue describir la categoría de la palabra a partir de su terminación morfológica. Esto es fundamental a la hora de trabajar con textos de aprendientes del español, en los que se deberá detectar y analizar las formas idiosincrásicas propias de la interlengua.

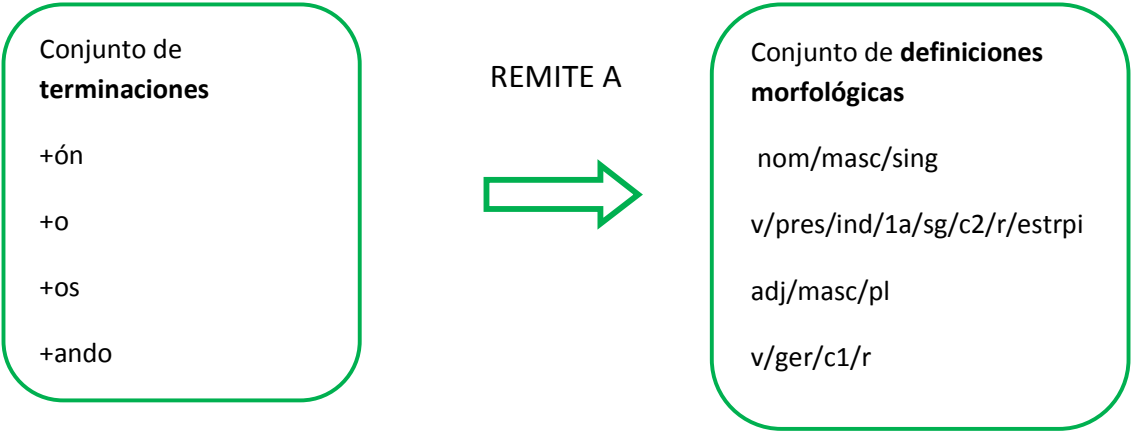
Cabe destacar que la información utilizada por Smorph está separada de la maquinaria algorítmica, por lo tanto permite adaptarla al uso que necesite darse, de modo tal que con el mismo software se puede tratar cualquier lengua e incluso la interlengua de aprendientes de español, si se cambia la información lingüística declarada en sus archivos.

Esta herramienta compila, minimiza y compacta la información lingüística de modo que quede disponible en un archivo binario. Los códigos fuente se dividen en cinco archivos:

- **Códigos Ascii:** en este archivo se especifican, entre otros, los caracteres separadores y las equivalencias entre mayúsculas y minúsculas.
- **Entradas:** funciona como un diccionario lingüístico. Aquí se ingresan los ítems léxicos que deben ir acompañados por un indicador del modelo correspondiente. Por ejemplo: corazón @n12. Este indicador de modelo es la conexión con el correspondiente archivo.
- **Terminaciones:** es fundamental declarar todas las terminaciones que son necesarias para definir los modelos de flexión. Dichos caracteres expresan un rasgo o conjunto de rasgos determinados. Por ejemplo, algunas terminaciones para nombres son: o, a, os, as, s, es, z, etc.

Si en la definición de un modelo se especifica una terminación no declarada en este archivo, el programa emite un mensaje de error. Las terminaciones se declaran una a continuación de otra, separadas por un punto. Es posible declarar una terminación vacía mediante el carácter @ y una terminación distinguida asociando a una terminación la definición morfológica correspondiente.

.



**Gráfico Nº 7 Terminaciones**

- En Rasgos se organizan jerárquicamente las etiquetas, por ejemplo, nombre, adjetivo, verbos, etc. Asimismo, se puede incorporar la etiqueta que indica, por ejemplo, el tipo de nombre, y se agregan los rasgos de concordancia, género y número.
- Modelos: En este archivo se introduce la información correspondiente a los modelos de flexiones morfológicas. Un modelo de flexión agrupa todas las flexiones de una misma clase de palabras. Esto se describe asociando a un conjunto de terminaciones el correspondiente conjunto de definiciones morfológicas. El esquema para definir los modelos es el siguiente:

<nombre\_modelo> -<cantidad de caracteres a sustraer>

<terminación 1> <definición morfológica para terminación 1>

Estos son introducidos mediante @, que indica el lugar en donde se ubica la raíz a la cual se agregan las terminaciones. Además, se consigna la información morfológica. Por ejemplo, para un modelo de sustantivo como niño:

@n4 -1

+o nom/masc/sg

+a nom/fem/sg

+os nom/masc/pl

+as nom/fem/pl

Para un sustantivo que se flexiona como la palabra futón, tenemos por ejemplo:

@n12 -2

+ón nom/masc/sg

+ones nom/masc/pl .

Aclaraciones: nom (nombre), masc (masculino), fem (femenino), sg (singular), pl (plural).

Se declara en primer lugar el nombre del modelo, indicado con el número, 4 y 12 en los ejemplos. Luego se declara la cantidad de caracteres que hay que sustraer a la forma lematizada. Este valor debe ser una cifra entre 0 y 9 y estar precedida del signo «-». Para formar el femenino de *niño*, por ejemplo, debe sustraerse el último carácter que es la *o*. En el caso de «futón» se quitan dos caracteres para poder eliminar la tilde que no se conserva en el plural. En tercer lugar, se ingresa la terminación, la cual debe estar declarada previamente en el archivo Terminaciones. La información morfológica corresponde a una cadena de caracteres sin espacios en blanco introducida por el signo + que indica adición al lema, por ejemplo: +ón.

Existe un archivo llamado **Terminaciones distinguidas** que contiene todas las terminaciones:

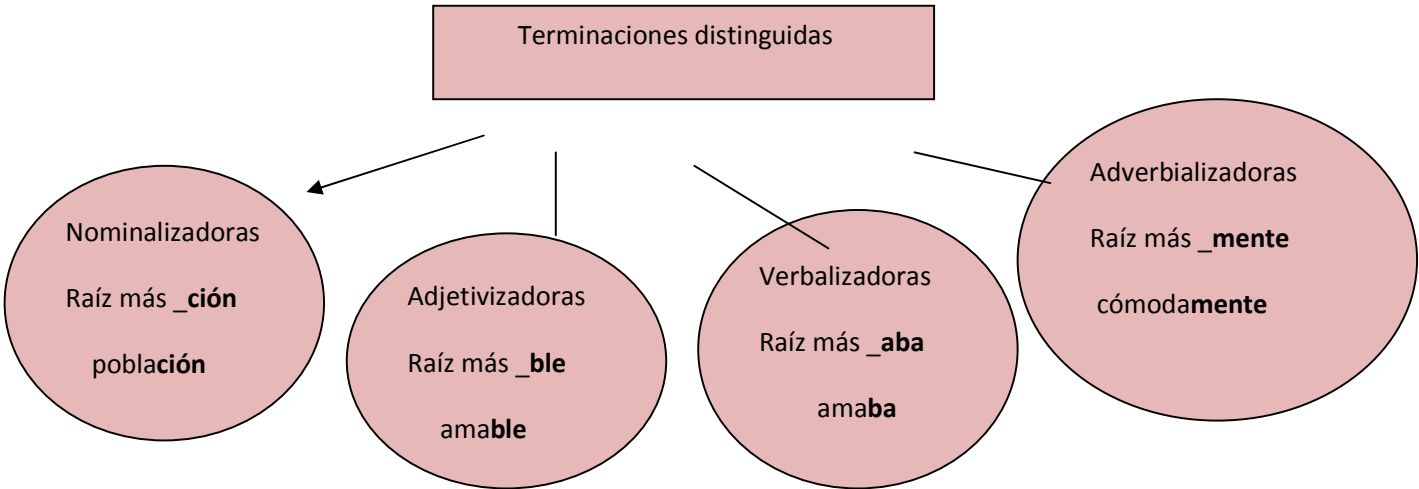
En los cuadros que siguen se presenta un fragmento del archivo:

endo .	iendo .
er .	ido .
ir .	as .
ar .	ás .
ando .	a .
ado .	amos .
abas .	an .
aba .	áis .
ábamos .	
aban .	

#### 8.1.1. Reconocimiento de palabras a partir de terminaciones distinguidas:

Las terminaciones distinguidas son los morfemas que se concatenan a la raíz de una palabra, asignando una categoría gramatical en particular. Por lo tanto, ante una palabra desconocida, es decir que no se encuentra consignada en el archivo *Entradas*, la terminación distinguida nos permitirá reconocer su categoría. De esta manera, cualquiera sea la cadena de caracteres acabada por *\_ción* (como por ejemplo *contracción*, *demolición*, *distinción*) podemos estar seguros de que dicho término es un nombre. Ait Mokhtar-Rodrigo Mateos (1995) presentan esta posibilidad de reconocimiento de palabras con el analizador SMORPH, atendiendo a su terminación, para solucionar el problema de los neologismos o de faltas en el diccionario.

Por ejemplo, en el caso de *demolición* *-ción* es un sufijo de nombre y en el caso de *contable* *-ble* es un sufijo de adjetivo. De acuerdo a la categoría gramatical que suministran a la raíz podemos clasificar a las terminaciones<sup>22</sup> en cuatro tipos:



**Gráfico N° 8: Terminaciones distinguidas**

Para que el programa reconozca las terminaciones distinguidas deberemos declarar en el archivo *Terminaciones* a estos sufijos junto con la información correspondiente:

<sup>22</sup> Solana Z., Tramallino C., Méndez B. y otros (2014) *Estudios del lenguaje: niveles de representación lingüística* (Proyecto Latin. Iniciativa latinoamericana de escritura de libros abiertos).

Ejemplo:

ción/nom

ble/adj

Cuando el reconocimiento automático ha sido hecho a partir de las terminaciones, el resultado arrojado será:

```
'temible'.  
[ 'temible', 'EMS','adj].
```

### 8.1.2. Análisis automático de Smorph

Para mostrar el análisis automático de Smorph empleo el fragmento de un texto de interlengua correspondiente al corpus inglés.

**Estoy bien porque viaje a las montañas con mi familia este semana. Fuemos al pueblo de Inlet Forge donde vi muchos animales.** Mi escuela termino la semana pasado. Muchos del dios voy al cine con mi novio o camino en el parque. Me gusta caminar mucho porque necesito el ejercicio. Tambien hago yoga con solamente mi porque la silencia esta muy tranquila. La moda es mi favorita campo de estudio. Tambien quiero estudiar en Australia porque me encanto viajar. Espero que hablamos pronto

En primer lugar copiamos y pegamos el texto escogido en el archivo texte.txt y luego de indicarle al programa que realice el análisis y lo muestre en este fichero obtenemos en el archivos smorpg.txt el siguiente resultado.

(Se muestra sólo el análisis de las dos primeras oraciones)

'Estoy'.

[ 'estar', 'EMS','v', 'MODOV','ind', 'PERS','1a', 'NUM','sg', 'TPO','pres'].

'bien'.

[ 'bien', 'EMS','adv'].

'porque'.

[ 'porque', 'EMS','cj'].

'viaje'.

[ 'viaje', 'EMS','nom', 'GEN','masc', 'NUM','sg'].

[ 'viajar', 'EMS','v', 'MODOV','subj', 'PERS','1a', 'NUM','sg', 'TPO','pres', 'TR','r', 'TC','c1', 'TDIAL','estrpi'].

[ 'viajar', 'EMS','v', 'MODOV','subj', 'PERS','3a', 'NUM','sg', 'TPO','pres', 'TR','r', 'TC','c1', 'TDIAL','estrpi'].

[ 'viajar', 'EMS','v', 'MODOV','imper', 'PERS','2a', 'NUM','sg', 'TPO','pres', 'TR','r', 'TC','c1', 'TDIAL','estrpi'].

'a'.

[ 'a', 'EMS','prep'].

'las'.

[ 'el', 'EMS','det', 'TDET','art'].

'montañas'.

[ 'montaña', 'EMS','nom', 'GEN','fem', 'NUM','pl'].

'con'.

[ 'con', 'EMS', 'prep' ].

'mi'.

[ 'mi', 'EMS', 'det', 'TDET', 'pos' ].

'familia'.

[ 'familia', 'EMS', 'nom', 'GEN', 'fem', 'NUM', 'sg' ].

'este'.

[ 'este', 'EMS', 'det', 'TDET', 'dem' ].

'semana'.

[ 'semana', 'EMS', 'nom', 'GEN', 'fem', 'NUM', 'sg' ].

'.'.

[ 'pf', 'EMS', 'pun' ].

'Fuemos'.

[ 'Fuemos', 'mi' ].

'al'.

[ 'al', 'EMS', 'contr' ].

'pueblo'.

[ 'pueblo', 'EMS', 'nom', 'GEN', 'masc', 'NUM', 'sg' ].

'de'.

[ 'de', 'EMS', 'prep' ].



'Inlet'.

[ 'Inlet', mi].

'Forge'.

[ 'Forge', mi].

'donde'.

[ 'donde', 'EMS', 'rel'].

'vi'.

[ 'ver', 'EMS', 'v', 'MODOV', 'ind', 'PERS', '1a', 'NUM', 'sg', 'TPO', 'prets', 'TR', 'hi', 'TDIAL', 'estrpi'].

'muchos'.

[ 'muchos', 'EMS', 'det', 'TINDF2', 'indf2a'].

'animales'.

[ 'animal', 'EMS', 'nom', 'GEN', 'masc', 'NUM', 'pl'].

Observaciones: Analiza como palabras desconocidas (m.i.) a *Inlet Forge* por ser un nombre propio en otra lengua que no está cargado en el diccionario y además a la forma verbal desviada *fuemos*. Con respecto al verbo *viaje* (viajé), al no llevar tilde el programa arroja todas las posibilidades: que es un nombre, que es un verbo en Subjuntivo o que pertenece al modo Imperativo.

El usuario puede emplearlo para analizar textos literarios, extraer información desde diarios o corpus técnicos, efectuar la formalización de fenómenos lingüísticos y analizar automáticamente textos.

Como se especifica en su manual (Silberztein, 2003), Nooj incluye herramientas para crear y mantener fuentes lexicales así como gramáticas sintácticas y morfológicas. Estas maquinarias, al encontrarse integradas, permiten que se desarrollen las operaciones morfológicas sobre los ítems mientras se realiza un análisis sintáctico. Un ejemplo de ello puede ser o controlar la concordancia morfosintáctica.

En cuanto a tareas de reconocimiento, los usuarios pueden desarrollar extractores para identificar unidades semánticas en textos extensos, tales como nombres de personas, ubicaciones, fechas, expresiones técnicas, entre otras.

Tanto para que el sistema reconozca por ejemplo, los verbos en un texto, como para que analice morfológicamente enunciados es necesario crear nuevos textos e ingresar los ítems léxicos, modelos de flexión y categorías con las que debe operar el software. Estos se consignan en tres archivos de diferente extensión: los modelos se ubican en el archivo Gramática («Grammar») que posee extensión .nog, los lemas se ingresan en el archivo Diccionario («Dictionary») con extensión .dic y, por último, las categorías gramaticales con sus rasgos, por ejemplo, Nombre: género/ número, en el archivo Propiedades («Properties ' definition»), que tiene extensión .def

Existen dos tipos de recursos lingüísticos:

- Dictionarios: (archivos con terminación .dic) asocian palabras o expresiones a por ejemplo una categoría, como verbo, a uno o más paradigmas de inflexiones o derivaciones (por ejemplo, cómo conjugar verbos), y también asocian a una o más propiedades sintácticas y semánticas.

Para cada idioma, Nooj tiene acceso a un diccionario en el que cada palabra de esa lengua es una entrada y se asocia a cierta información morfológica, por lo general a su inflexión o paradigmas derivativos. El paradigma flexivo dice a Nooj cómo flexiona la entrada léxica, es decir, cuáles son sus formas conjugadas (si se trata de un verbo) y su femenino y plural (para los nombres en las lenguas románicas).

El siguiente cuadro es un fragmento del diccionario correspondiente a verbos. Se indica que use la gramática inflexional llamada verbos y luego se ingresa cada verbo en infinitivo

con la indicación del modelo de inflexión que debe seguir, por ejemplo acabar es igual a la conjugación de amar:

```
#use verbos.nof  
abrir,V+FLX=ABRIR  
acabar,V+FLX=AMAR  
comer,V+FLX=TEMER  
crecer,V+FLX=CONOCER  
criar,V+FLX=ACENTUAR  
cubrir,V+FLX=CUBRIR  
cumplir,V+FLX=PARTIR  
dar,V+FLX=DAR  
deber,V+FLX=TEMER  
decir,V+FLX=DECIR  
dedicar,V+FLX=BUSCAR  
degollar,V+FLX=ACERTAR  
dejar,V+FLX=AMAR  
demostrar,V+FLX=ACERTAR
```

Gramáticas: se usan para representar fenómenos lingüísticos, desde niveles ortográficos y morfológicos a niveles sintagmáticos. Contiene tres tipos de gramáticas:

- Gramáticas derivacionales o (archivos con terminación .nof), que se usan para representar las propiedades flexivas, por ejemplo, conjugaciones o derivaciones como las nominalizaciones de las entradas léxicas. Estos modelos pueden declararse de forma gráfica o mediante reglas.
- Gramáticas lexicales, ortográficas, morfológicas o terminológicas (archivos con terminación .nom). Se usan para representar conjuntos de tipos de palabras y asociarlas con información léxica; por ejemplo: estandarizar la ortografía de palabras o de variantes, reconocer neologismos, asociar expresiones sinónimas.
- Gramáticas semánticas o sintácticas (archivos con terminación .nog). Son utilizadas para reconocer o anotar expresiones en textos como, por ejemplo, etiquetar tanto a frases nominales como a ciertas expresiones sintácticas o idiomáticas. Además, se emplean para extraer o desambiguar palabras filtrando algunos ítems léxicos o anotaciones sintácticas en el texto.

Los paradigmas inflexión-derivación son formalizados como bibliotecas de gráficos (grafos) estructurados o reglas basadas en textos.

La siguiente captura de pantalla muestra la gramática sintáctica referida a frases verbales pasivas formadas con el verbo ser más el participio:

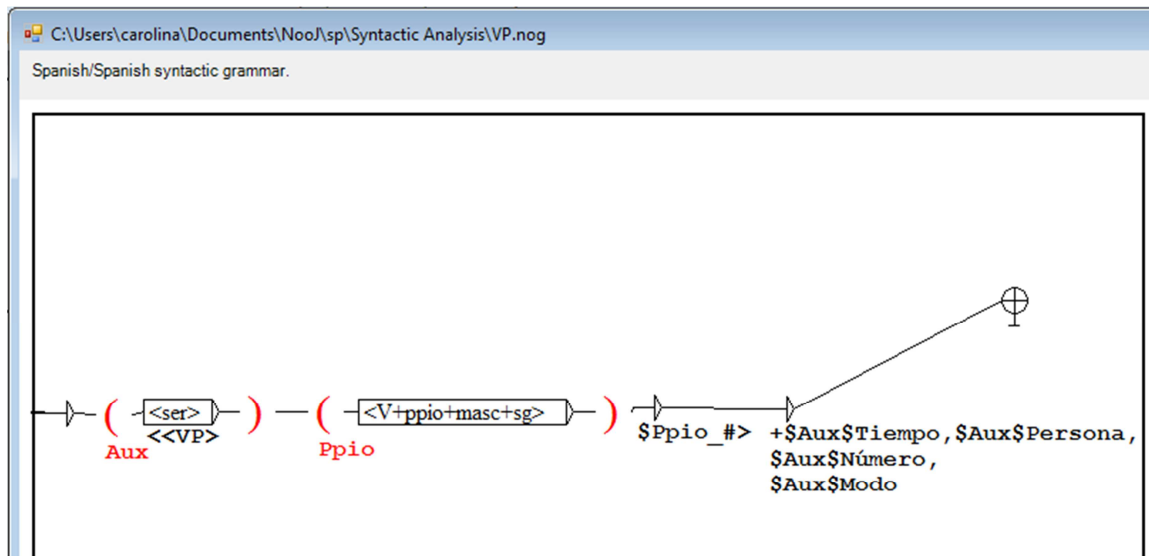


Imagen de captura de pantalla N° 1: ejemplo de gramática sintáctica para frases verbales.

La captura de pantalla N° 2 muestra las gramáticas sintácticas para reconocer sintagmas nominales en un texto:

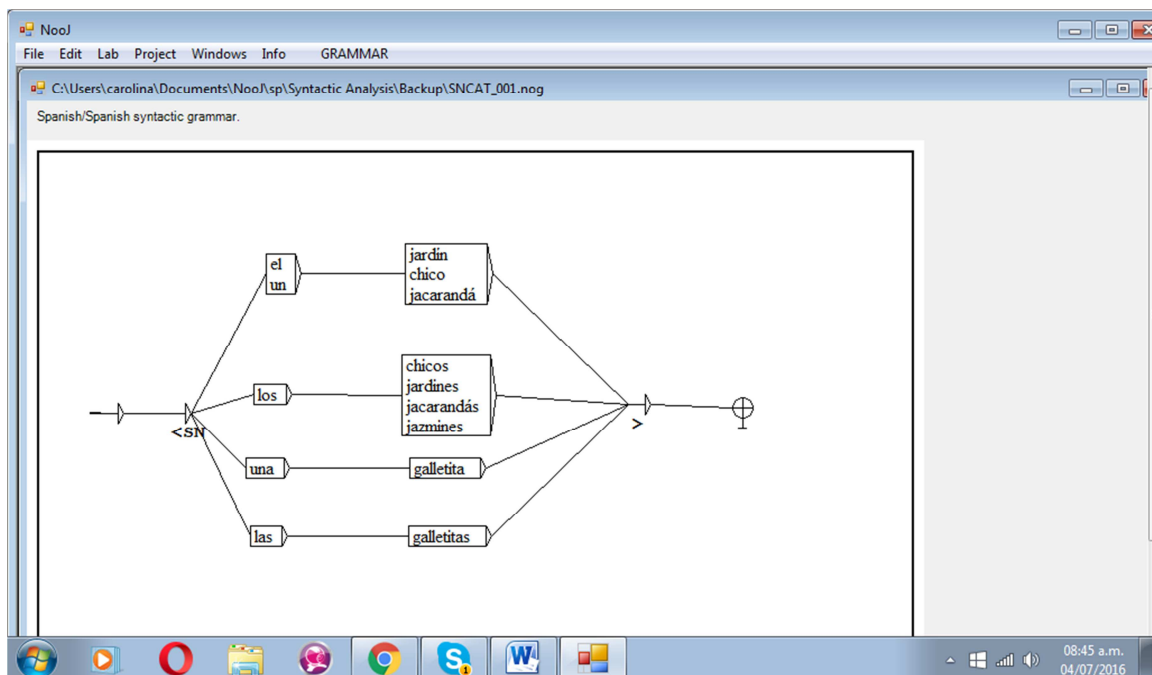


Imagen de captura de pantalla N° 2: ejemplo de gramática sintáctica para sintagmas.

### 8.2.1. Análisis morfológico de Nooj

Página | 118

Como muestra voy a analizar un texto de interlengua perteneciente al corpus alemán:

*Ustedes todos saben de mi situación. Estoy en 5to y no tengo idea que quiero estudiar. Ayer hablé con mis padres, ellos me aconsejaron que estudiara política porque siempre me sacé buenas notas pero yo les dije que no haya querido ser político.*

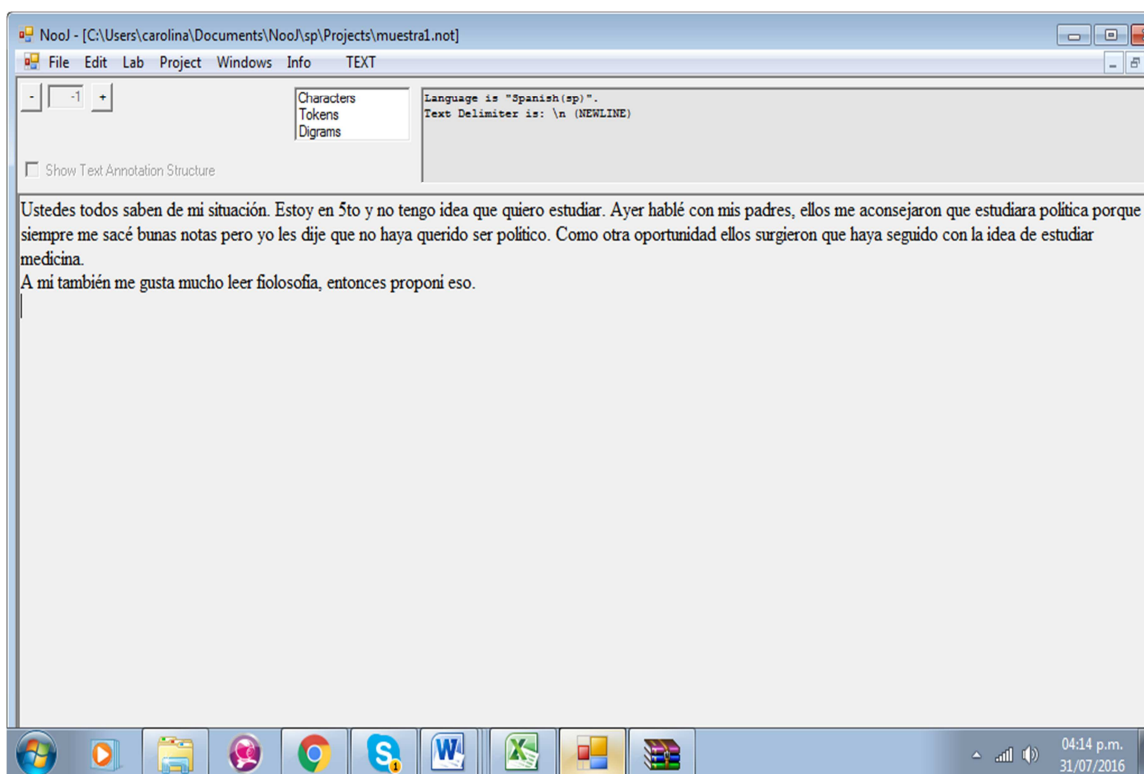
*Como otra oportunidad ellos surgieron que haya seguido con la idea de estudiar medicina.*

*A mí también me gusta mucho leer la filosofía, entonces proponí eso.*

El primer paso es generar este texto en el programa. Para eso, en primer lugar, hacemos clic en File > New > Text y verificamos que en la nueva ventana se encuentre seleccionada la opción sp, correspondiente al módulo de español.

Luego de escribirlo o en este caso de copiarlo y pegarlo en la ventana que se ha abierto, hacemos clic en File > Save, y le asignamos un nombre, voy a llamarlo MUESTRA1 (el programa agrega automáticamente la extensión .not) y lo guardamos tocando el botón Guardar.

Luego hacemos clic en TEXT y dentro de esta ventana en Linguistic Analysis y obtenemos la siguiente ventana:



### Imagen captura de pantalla N° 3

A continuación, podemos observar el análisis automático que produce NooJ:

Inmediatamente debajo de la barra de menús y por encima del texto, se puede ver – de izquierda a derecha– la siguiente información:

1. Un recuadro con un número (1, en este caso) precedido por el signo – y seguido por +/3TUs:
  - a) 3TUs indica el número de unidades textuales del archivo; se refiere a los tres párrafos que escribimos (los párrafos coinciden con oraciones, los párrafos con más de una oración cuentan como una sola unidad textual), si modificáramos el texto agregando o quitando oraciones, ese número se modificaría.
  - b) El número indica cuál es la unidad textual que se está analizando en un momento determinado. Como el texto no tiene anotaciones, por ahora, esta información es irrelevante.
2. Debajo de ese recuadro, se lee Show Text Annotation Structure precedido de un espacio para seleccionar esa opción. Cuando agreguemos diccionarios y gramáticas, nos permitirá ver el resultado de su aplicación en el texto.

3. En la parte central vemos una pequeña ventana con una barra de desplazamiento. Cada una de las líneas de esa ventana, que se pueden seleccionar haciendo doble clic sobre ellas, da alguna información sobre el texto:

a) Characters: indica la frecuencia con la que aparece cada carácter en el texto, el tipo de carácter –los caracteres pueden ser letras, espacios en blanco o delimitadores (signos de puntuación)– y el código numérico que le corresponde a cada uno de ellos en el sistema Unicode. Haciendo clic en la barra superior se pueden seleccionar distintos criterios para ordenarlos; por ejemplo, si hacemos clic en Freq los ordenará de mayor a menor por la frecuencia con que aparecen en el texto, en nuestro caso el carácter que aparece con mayor frecuencia es el espacio en blanco, que aparece cuarenta y dos veces.

b) Tokens: indica la frecuencia con la que aparecen conjuntos de caracteres delimitados por espacios en blanco. Como se puede observar, las variaciones morfológicas de cada palabra constituyen distintos tokens. Aquí marca 78 tokens.

Además indica que hay 70 palabras y 7 delimitadores (que son los puntos).

c) Digrams: marca las agrupaciones de dos tokens que se repiten en el texto. En nuestro caso, la única secuencia que se repite dos veces es los chicos.

d) Annotations: abre automáticamente un diccionario en blanco para que se incorporen anotaciones al texto.

e) Unknowns: marca las palabras que no aparecen en ningún diccionario. En este caso, las selecciona a todas porque no hay diccionarios.

f) Ambiguities: cuando hay diccionarios, indica cuáles son las palabras que tienen más de una entrada en los diccionarios. Por ejemplo, cuando puede estar etiquetado como sustantivo y como verbo.

g) Unambiguous Words: marca las palabras que reciben una sola etiqueta en los diccionarios.

Ahora bien, para que realice el análisis morfológico debemos indicarle con qué diccionarios (que a la vez están asociados a gramáticas morfológicas) debe analizar el texto. Estos diccionarios deben compilarse previamente. Para ello hacemos clic en Lab> dictionary y con el botón set seleccionamos el diccionario.

Para aplicar el diccionario al análisis del texto, debemos abrirlo haciendo clic en File > Open > Text (el texto se puede modificar haciendo clic en TEXT > Modify Text). Luego hacemos clic en Info > Preferences; en la pestaña General, verificamos que esté



seleccionado el idioma sp; en la pestaña Lexical Analysis, tocamos el botón Refresh para que aparezca el diccionario que hemos compilado. Una vez que hayamos seleccionado todos los archivos que queremos aplicar al análisis, hacemos clic en *Apply*.

Para efectuar el análisis, hacemos clic en *TEXT > Linguistic Analysis* y verificamos los resultados.

Para efectuar el análisis, hacemos clic en *TEXT > Linguistic Analysis* y verificamos los resultados.

Tildando el casillero que está a la izquierda de Show Text Annotation Structure, en la parte inferior de la ventana aparece un análisis de la primera oración, cada vez que hacemos clic en “+” pasamos a la siguiente oración y vemos el análisis de cada una de ellas.

Las siguientes capturas de pantalla muestran el análisis de la última oración del texto de interlengua:

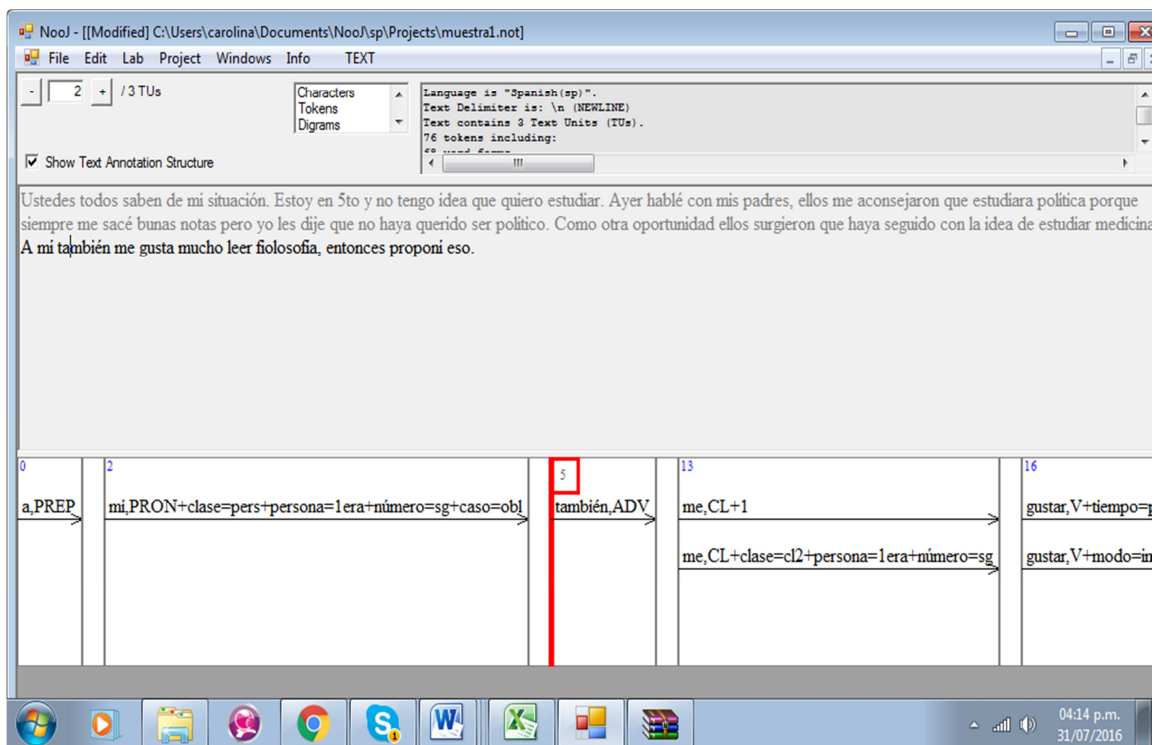
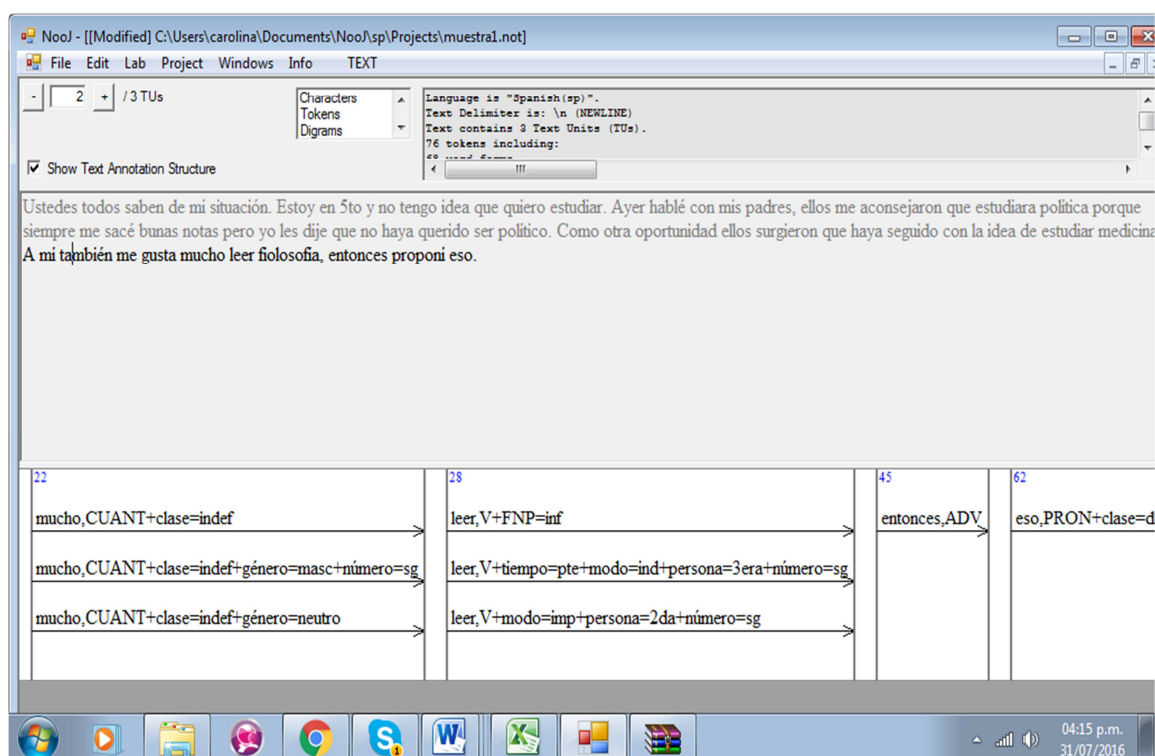


Imagen captura de pantalla N° 4



**Imagen captura de pantalla N° 5**

Como vemos no analiza las palabras propias de la interlengua: *filosofía* y *proponí*, por no hallarse en los diccionarios de español.

### 8.2.2. Localización de términos en NooJ

Para localizar palabras hacemos clic en TEXT > Locate y escribimos cualquier token o secuencia de tokens del texto, tocamos cualquiera de los botones que se encuentran en la parte inferior de la pantalla y NooJ muestra todas las veces que aparece en el texto; por ejemplo, si buscamos la palabra *estudiar*, obtenemos el siguiente resultado:

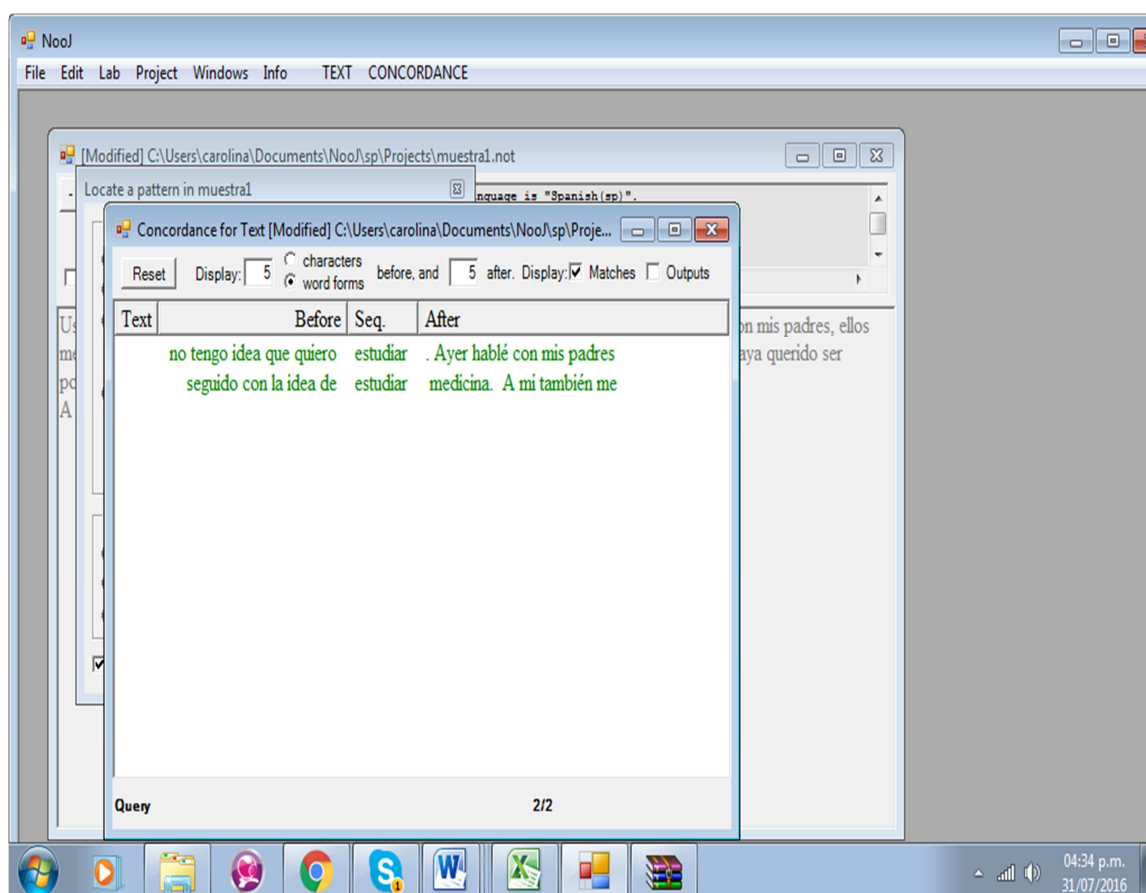


Imagen captura de pantalla N° 6

En la parte superior de la pantalla se pueden seleccionar la cantidad de palabras o caracteres del contexto anterior o posterior que queremos ver (en este caso cinco). En la parte inferior derecha 2/2 indica que en la pantalla se muestran dos coincidencias con la búsqueda, que es el total de las coincidencias.

Además de buscar una categoría con un rasgo morfológico determinado, se pueden buscar todas las variantes morfológicas de un lema, para hacerlo debemos escribir en minúsculas cualquier variante del lema entre <>. Por ejemplo para buscar todas las formas flexionadas de un verbo ingresamos el verbo en infinitivo. Probamos con el verbo tener y el resultado de la búsqueda es:

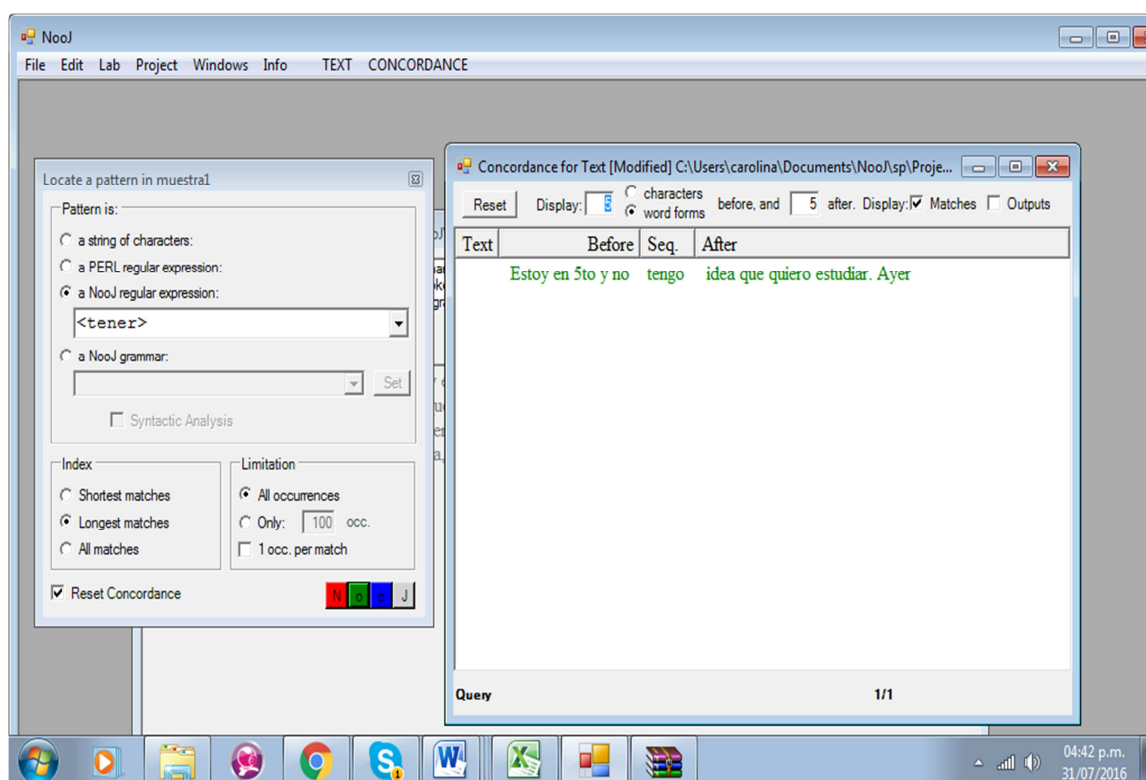


Imagen captura de pantalla N° 7

Lo interesante es que podemos indicarle que busque todas las ocurrencias o sólo 100 de ellas, esto es fundamental para trabajar con datos estadísticos en grandes corpus de textos.

## CAPÍTULO 9: TRATAMIENTO AUTOMÁTICO PARA LOS VERBOS IRREGULARES DEL ESPAÑOL ESTÁNDAR

En este capítulo se explica implantación en máquina de la morfología verbal para el análisis automático de ambas herramientas informáticas: Smorph y NooJ. En cuanto a este último, aún nuestro equipo de trabajo se encuentra completando la cantidad de gramáticas necesarias para contemplar todas las irregularidades que presenta la lengua española empleando el mismo criterio de agrupamiento y sistematización utilizado para Smorph.

## 9.1. Modelos verbales irregulares de Smorph

Los rasgos que distinguen la irregularidad para los modelos verbales irregulares son:

Rasgos vocálicos (diptongación o cierre de vocales), consonánticos (cambio o introducción de una consonante), cambios de acentuación, cambio total (o casi total) de raíz y cambio de desinencias respecto de las del verbo tipo.

**vo/ie**    vocálico, cambia 'e' por 'ie', ejemplo: *acert-ar /aciert-o*

**vo/ ue**    vocálico, cambia 'o' por 'ue', ejemplo: *vol-ar/vuel-o*

**vo/i**      vocálico, cierra 'e' de la raíz en 'i', ejemplo: *ped-ir /pid- o*

**vo/u**      vocálico, cierra 'o' de la raíz en 'u', ejemplo: *dorm-ir/durm-ió*

**vo/y**      vocálico, transforma la vocal 'i' en semivocal 'y' *aflu-ir/ afluyo*

**ac**        cambia acentuación, ejemplo: *actu-ar /actú-o*

**co/ggu**    consonántico, raíz terminada en 'c' cambia por 'qu', ejemplo: *sac-ar/saqu-é*

**co/gj**      consonántico, raíz terminada en 'g' cambia por 'j', ejemplo: *dirig-ir /dirij-o*

**co/cz**      consonántico, raíz terminada en 'c' cambia por 'z', ejemplo: *zurc-ir/zurz-o*

**co/zc**      consonántico, raíz terminada en 'z' cambia por 'c', ejemplo: *comenz-ar/comenc-é*

**co/lg**      consonántico, inserción de 'g' después de la 'l' final de la raíz, ejemplo: *val-er/valg-o*

**co/fj**      consonántico con infijjo, ejemplo: *permanec-er/permanezc-o*

**co/que**    consonántico, raíz terminada en 'qu' cambia en 'c', ejemplo: *Delinqu-ir/delinc-o*

**voco**      cambio consonántico y vocálico a la vez

**cr** cambio de raíz

En el archivo Rasgos los encontramos consignados de la siguiente manera:

**MODOV    modover**

ind    indic

subj    subjun

imper    imperat

infin    infinit

infcl    infinit+cl

ger    gerun

dsvger    desvo

part    parti

dsvprt    desvpa .

**NUM            numero**

sg    sing

pl    plur

**PERS        perso**

1a    prim

2a    segun

3a    terc .

**TC        tconj**

C1    conj1

C2    conj2

C3    conj3

**TR        treg**

r    reg

ir    irreg

irr    irre

hi    hiper .

**TIRR        tirreg**

vo    voc

co    cons

voco    vocons

ac    acen

acco    acenc

sinc    sincop

cr    craiz

des    desin .

Para tratar de reunirlos en el menor número posible de tipos, con el grupo INFOSUR factorizamos los verbos irregulares atendiendo a las personas y tiempos en que la irregularidad tiene lugar, por eso hay casos en que es desatendido el tipo de irregularidad que experimenta la raíz, priorizando la economía del sistema por sobre la descripción ya que nuestro punto de vista está determinado por las necesidades del trabajo en máquina. Así quedan reunidos en el mismo modelo:

acert-ar      acier-to

son-ar      suen-o

acentu-ar      acentú-o

mec-er      mez-o

val-er      valg-o

luc-ir      luzc-o

dirig-ir      dirij-o

Dicho de otra manera, se agrupan los verbos irregulares según compartan las terminaciones verbales, por lo tanto siguen un mismo modelo los verbos terminados en ar que diptongan su raíz, así cambie o por el diptongo /ue/ o bien e por el diptongo /ie/.

### 9.1.1. Ejemplos de modelos verbales irregulares

Acertar      aciert v4/vo/ie

Acertar      v5/fr 5

Se conjugan como 'acostar', 'aprobar', 'calentar', 'cerrar', 'despertar', 'encerrar', 'gobernar', 'mostrar', 'pensar', 'quebrar', 'recordar', 'sembrar, etc.

jugar	jueg	v6/vo/ue
jugar	juegu	v7/voco/ue/ggu
jugar	jugu	v8/co/ggu
jugar		v9/ fr9

Se conjugan como 'almorzar', 'calentar', 'cerrar', 'despertar', 'encerrar', 'gobernar', 'mostrar', 'pensar', 'quebrar', 'recordar', 'sembrar?', etc.

Para el verbo andar en el que cambian las terminaciones tenemos dos modelos, uno para las raíces regular y otro para las irregulares.

Andar	and	v12/cr
andar	anduv	v14/cr

El siguiente cuadro muestra la conjugación del verbo andar en el Modo Indicativo, destacando las desinencias que varían:

Presente	Pret.Imp.	Pret.Simp.	Fut.Imp.	Cond
1a.sg. and-o	and-aba	and- <b>uve</b>	and-aré	and-aría
2a.sg. and-as	and-abas	and- <b>uviste</b>	and-arás	and-arías
2a.sg. and-ás	and-abas	and- <b>uviste</b>	and-arás	and-arías
3a.sg. and-a	and-aba	and- <b>uvo</b>	and-ará	and-aría
1a.pl.and-amos	and-ábamos	and- <b>uvimos</b>	and-aremos	and-aríamos
2a.pl. and-áis	and-abais	and- <b>uvisteis</b>	and-aréis	and-aríais
2a.pl. and-an	and-aban	and- <b>uvieron</b>	and-arán	and-arían
3a.pl. and-an	and-aban	and- <b>uvieron</b>	and-arán	and-arían

**Tabla Nº 52 : conjugación del verbo andar**

### 9.1.2. Verbos con cambios gráficos

Con excepción de 'dar' y 'andar', que tienen irregularidades complejas, los verbos de la primera conjugación presentan el siguiente esquema:



La diptongación y los cambios de acentuación se producen solamente en algunos verbos con raíces cuya última vocal no es -a y en las formas en que la raíz es tónica, lo que implica que la terminación es átona. Dicho de otro modo, con desinencias átonas las raíces terminadas en -e, -i, -o, -u pueden ser irregulares, ejemplo: acertar, acierto.

Los verbos terminados en -gar, -guar -zar y -car, modificarán la consonante en casos en que la terminación comience con -e, -i :

Para los verbos con cambios gráficos de la primera conjugación hallamos:

Terminación	Cambia	Por	ejemplo
<b>-car</b>	-c	-qu	Tocar-toqué
<b>-zar</b>	-z	-c	Comenzar-comencé
<b>-gar</b>	-g	-gu	Tragar-tragué
<b>-guar</b>	-gu	-gü	Averiguar-averigüé

**Tabla Nº 53: verbos con cambios gráficos**

Algunos verbos manifiestan una combinación de ambas irregularidades, que podemos representar en el siguiente esquema:

1º segmento	1ª vocal de la terminación	Cambio	Ejemplo
átono	-a, -o	sin cambios	jugamos, jugó
átono	-e	cambio de la consonante	jugué
tónico	-a, -o	diptongación	juega, juego
tónico	-e	diptongación y cambio de la consonante	juegue

**Tabla Nº 54: ejemplo verbo con cambio gráfico**

En la segunda conjugación encontramos mayor número de verbos que presentan irregularidades complejas o no sistemáticas, no obstante, para muchos resulta válido el esquema que proponemos para la primera, con algunas modificaciones.

La diptongación y los cambios acentuales y vocálicos también afectan solo a las raíces tónicas. Ejemplo: tender / tiendo.

Los cambios consonánticos afectan a todas las formas cuyas terminaciones comienzan en -a, -o:

Terminación	Cambia	Por	<i>Ejemplo</i>
-cer	-c	-z /-zc	ejercer / ejerzo; parecer / parezco
-ger	-g	-j	proteger / protejo

**Tabla Nº 55: cambios consonánticos**

La combinación de ambas irregularidades se produce según el siguiente esquema:

1º segmento	1ª vocal de la terminación	Cambio	Ejemplo
átono	-e, -i	sin cambios	cocí, cocemos
átono	-a	cambio consonante	cozamos
tónico	-e,	diptongación	cuece
tónico	-a, -o	diptongación y cambio consonante	cueza, cuezo

**Tabla Nº 56: cambios consonánticos**

### 9.1.3. Verbos hiperirregulares

Finalmente, existe un grupo de verbos que llamamos hiperirregulares, puesto que presentan una flexión particular, con irregularidades propias. En la primera conjugación,

se incluyen en esta clasificación los verbos *estar*, *dar* y *andar*, que son regulares en la raíz, pero presentan desinencias especiales.

Los criterios que hemos usado para agrupar estos verbos son:

- El caso 1: a) Verbos en -er con cambio vocálico en 1a.pers.sg, 2a.pers.sg, 3a.pers.sg y 3a.pers.pl.: **puedo, quiero.**
- b) Verbos en -er con cambio consonántico1 (agregado) en 1a.pers.sg: **pongo, tengo (pon-g-o, ten-g-o)**
- c) Verbos en -er con cambio vocálico y consonántico1 (agregado) en 1a.pers.sg: **traigo, caigo (traí-g-o, caí-g-o)**
- d) Verbos en -er con cambio consonántico2 en 1a.pers.sg: **hago. (hac> hag).**
- e) Verbos en -er con cambio vocálico y consonántico2 en 1a.pers.sg: **quepo (cab>quep).**
- f) Verbos en -er con otros cambios en 1a.pers.sg: **sé.**
- g) Verbos en -ir con cambio vocálico y alternancia de consonante en 1a.pers.sg, 2a.pers.sg, 3a.pers.sg y 3a.pers.pl.: **sigo (segu>sigu segu>sig).**
- h) Verbos en -ir con cambio consonántico1 (agregado) en 1a.pers.sg: **salgo, oigo, vengo (sal-g-o, oi-g-o, ven-g-o).**
- i) Verbos en -ir con cambio consonántico3(infijo) 1a.pers.sg: **conduzco**
- j) Verbos en -ir con cambio vocálico y consonántico1 en 1a.pers.sg: **digo (dec>dig).**

**caso 1 -er**

**caso 2 -er**

verbo	persona pres.	voc.	cons 1	cons 2	cons 3	verbo	pers pret.s.	des. reg	des. irr	voc	cons
<b>poder</b>	<b>1s/2s/ 3s /3pl</b>	<b>x</b>				<b>poder</b>	<b>todas</b>		<b>x</b>	<b>x</b>	

poner	1a.sg.		x			poner	todas		x	x	x
traer	1a.sg.	x	x			traer	todas		x		x
hacer	1a.sg.			x		hacer	todas 3a.sg. diferente		x	x	x
caber	1a.sg.	x		x		caber	todas		x	x	x
caso 1-- ir											
se guir	1s/2s/ 3s /3pl	x		x		seguir	3a.sg/ 3a.pl.	x		x	
salir	1a.sg.		x			salir	regula r				
con ducir	1a.sg.				x	conducir	todas		x		x
decir	1a.sg.	x	x			decir	todas		x	x	x

**Tabla Nº 57 verbos hiperirregulares**

El caso 2:

- Verbos en -er con desinencias regulares y con cambio consonántico en la raíz: **caí** (cayó).
- Verbos en -er con desinencias irregulares y con cambio vocálico en la raíz: **pude**.
- Verbos en -er con desinencias irregulares y con cambio vocálico y consonántico en la raíz: **pude, quise, puse, cupe, supe, hice, tuve**.
- Verbos en -er con desinencias irregulares y con cambio consonántico en la raíz: **traje**
- Verbos en -ir con desinencias regulares y sin ningún cambio en la raíz: **salí**
- Verbos en -ir con desinencias regulares con cambio vocálico en 3a.pers.sg. y 3a.pers.pl: **seguí (siguió)**.

g) Verbos en -ir con desinencias regulares con cambio consonántico en 3a.pers.sg. y 3a.pers. pl: **oí (oyó)**.

h) Verbos en -ir con desinencias irregulares y con cambio vocálico en la raíz: **vine**.

Página | 133

i) Verbos en -ir con desinencias irregulares y cambio consonántico en la raíz: **conduje, dije**.

Para el verbo *querer*, etiquetado como hiperrregular, perteneciente a la segunda conjugación, encontramos las formas flexivas irregulares indicadas de la siguiente manera en el archivo Entradas:

querer quiero /v/pres/ind/1a/sg/hi/estrpi .

querer quieres /v/pres/ind/2a/sg/hi/est .

querer quiere /v/pres/ind/3a/sg/hi/estrpi .

querer quieren /v/pres/ind/3a/pl/hi/estrpi .

querer quiera /v/pres/subj/1a/sg/hi/estrpi .

querer quieras /v/pres/subj/2a/sg/hi/estrpi .

querer quiera /v/pres/subj/3a/sg/hi/estrpi .

querer quieran /v/pres/subj/3a/pl/hi/estrpi .

querer quiera /v/pres/imper/2a/sg/hi/estrpi .

querer quiere /v/pres/imper/2a/sg/hi/est .

querer quieran /v/pres/imper/3a/pl/hi/estrpi .

querer quis @v36 .

querer quise /v/prets/ind/1a/sg/hi/estrpi .

querer quisiste /v/prets/ind/2a/sg/hi/estrpi .

querer quiso /v/prets/ind/3a/sg/hi/estrpi .

querer quisimos /v/prets/ind/1a/pl/hi/estrpi .

querer quisisteis /v/prets/ind/2a/pl/hi/est .

querer quisieron /v/prets/ind/3a/pl/hi/estrpi .

Para las formas flexivas que no varían hallamos la indicación a los siguientes modelos:

querer quer @v38 .

querer @v40 .

## 9.2. Modelos verbales en el sistema Nooj

Para el tratamiento de la morfología y la creación de diccionarios se utilizan tres archivos:

- El archivo Propiedades

En este archivo deben consignarse en mayúscula los rasgos que se utilizan para etiquetar las entradas del diccionario con sus respectivos valores que se declaran en minúscula. Como categorías se ingresan: VERBO, NOMBRE, ADJETIVO, ADVERBIO, PREPOSICIÓN, DETERMINANTE, CUANTIFICADOR, CONJUNCIÓN, CONTRACCIÓN, INTERJECCIÓN, CLÍTICO, PRONOMBRE, ARTÍCULO, etc.

El siguiente cuadro muestra las propiedades definidas:

```
CATEGORÍAS = V | N | ADJ | ADV | PREP | DET | CUANT | CONJ | CONTR | INTERJ | CL |
PRON | ART | REL | INTEXC | LOCLAT;

FLEXIÓN = inf | ppio | ger | pte | pi | pps | fut | cond | ind | subj | imp | inf
| ppio | ger | 1era | 2da | 3era | sg | pl | n_inv | masc | fem | neutro | g_inv | TC | ind | subj |
imp | 1era | 2da | 3era
| sg | pl | masc | fem;

#| apócope | cl1 | cl2 | cl3 | cl4 | PR | artdet | artindet | dem | pos | nom | obl;

DERIVACIÓN = dim | nom | neg;

V_FNP = inf | ppio | ger;

V_tiempo = pte | pi | pps | fut | cond | TC;

V_modos = ind | subj | imp;

V_persona = 1era | 2da | 3era;

V_género = masc | fem;

V_dialecto = RIOP | LUNF;
```

= ind | subj | imp;) Las personas: 1era, 2da y 3 era y el número singular o plural (V\_persona = 1era | 2da | 3era; V\_número = sg | pl; )

#### b. Gramática

Página | 135

En Gramática deben declararse los modelos de los verbos: se definen los paradigmas de flexión, estos archivos admiten cualquier denominación con extensión .nof, asignada automáticamente por el programa diccionario con terminación (.dic)

Para un verbo de la segunda conjugación regular, como por ejemplo *comer*, tenemos las siguientes terminaciones que se asocian a la raíz verbal:

**#amar, temer, partir**

**TER1 = (é/fut+ind+1era+sg | ás/fut+ind+2da+sg | á/fut+ind+3era+sg |  
emos/fut+ind+1era+pl | éis/fut+ind+2da+pl**

**| án/fut+ind+3era+pl | ía/cond+ind+1era+sg | ías/cond+ind+2da+sg |  
ía/cond+ind+3era+sg | íamos/cond+ind+1era+pl**

**| íais/cond+ind+2da+pl | ían/cond+ind+3era+pl);**

**#am, tem-i, part-i,**

**TER2 = (ó/pps+ind+3era+sg);**

#### c. Diccionario

Los verbos irregulares que diptongan como es el caso de *colar* y *comenzar* siguen el modelo de otros verbos irregulares: *acertar* y *almorzar*, respectivamente. A continuación se exhibe un fragmento del diccionario de verbos en donde se observa el paradigma de flexión que sigue cada verbo:

```
#use verbos.nof  
  
abrir,V+FLX=ABRIR  
acabar,V+FLX=AMAR  
acentuar,V+FLX=ACENTUAR  
acertar,V+FLX=ACERTAR  
ayudar,V+FLX=AMAR  
beber,V+FLX=TEMER  
colar,V+FLX=ACERTAR  
comenzar,V+FLX=ALMORZAR
```

En la primera línea del archivo se indica, antecedido del signo numeral, el nombre de la gramática que debe utilizar el programa; en las líneas siguientes, cada entrada seguida de coma con su etiqueta categorial y, en mayúscula, el paradigma de flexión que corresponda. Por ejemplo, [acabar, V+FLX=AMAR] indica que acabar corresponde a la categoría verbo y que se flexiona como amar. La denominación del paradigma de flexión es elegida por el usuario, es decir, se asigna de forma arbitraria en la gramática; en este ejemplo se utiliza AMAR.



## CAPÍTULO 10: ANÁLISIS AUTOMÁTICO DE LA INTERLENGUA

En este capítulo muestro las adaptaciones que debo realizar en los archivos de los programas informáticos con los que trabajo para poder lograr el reconocimiento automático de las formas verbales irregulares desviadas. En primer término explico las modificaciones en los archivos Entradas, Rasgos y Modelos de Smorph para el análisis automático de desv1, desv2 y desv3 y luego la creación de gramáticas productivas para detectar los cuatro casos de desviaciones de interlengua con el sistema Nooj. Por último presento una conclusión de los alcances obtenidos por una y otra herramienta.

### 10.1. Implementación en Smorph

A partir de los siguientes fragmentos de textos de aprendientes de español, obtenemos el análisis automático morfológico que realiza Smorph:

*Ejemplos Nº 1:*

*“Somos dos canadienses y queremos pasar una fin de semana en Rosario.*

*Puede escribir me como podemos reservar? ”*

*“Ellos quieren visitar la ciudad de Rosario.”*

*“Usted puede, por favor,irme la disponibilidad del apart.”*

*“Me gusta mucho Rosario y me gusta mucho todo que hacemos en nos viaje.*

*Ahora, yo tengo dos amigas canadienses que querian ir en vacaciones en la Argentina y van a pasar por rosario en el último fin de semana de septiembre.”*

*“Un museo donde se demuestran como se hacen los suecos y como se moleron el trigo en un molino.”*

*“También conta este artículo que el asesino y víctima hacían y la fecha del crimen.”*

*“Los policías deteneron al asesino en Mendoza, 21 de septiembre 2006...”*

*“Cuando hemos haido la fiesta el 10 de junio, vos haces una fotografia de nosotros.*

*“Después escuchamos que no tenemos habitación.”*

### **Análisis morfológico:**

Página | 138

'Somos'.

[ 'ser', 'EMS','v', 'MODOV','ind', 'PERS','1a', 'NUM','pl', 'TPO','pres'].

'dos'. [ 'num', 'EMS','cardin'].

'canadienses'.

[ 'canadiense', 'EMS','adj', 'GEN','\_', 'NUM','pl'].

'y'.

[ 'y', 'EMS','cop'].

'queremos'.

[ 'queremos', mi].

'pasar'.

[ 'pasar', 'EMS','v', 'MODOV','infin', 'TR','r', 'TC','c1'].

'una'.

[ 'una', 'EMS','det', 'TINDF2','indf2a'].

'fin'.

[ 'fin', 'EMS','nom', 'GEN','masc', 'NUM','sg'].

'de'.

[ 'de', 'EMS','prep'].

'semana'.

[ 'semana', 'EMS','nom', 'GEN','fem', 'NUM','sg'].

'en'.

[ 'en', 'EMS','prep'].

'Rosario'.

[ 'rosario', 'EMS','nom', 'GEN','masc', 'NUM','sg'].

'Puede'.

[ 'poder', 'EMS','v', 'MODOV','imper', 'PERS','2a', 'NUM','sg', 'TPO','pres', 'TR','hi'].

'escribir'.

[ 'escribir', 'EMS','v', 'MODOV','infin', 'TR','r', 'TC','c3'].

'me'.

[ 'lo', 'EMS','cl', 'TPCRF','rflnse'].

'como'.

[ 'como', 'EMS','rel'].

'podemos'.

[ 'podemos', mi].

'reservar'.

[ 'reservar', 'EMS','v', 'MODOV','infin', 'TR','r', 'TC','c1'].

'quieren'.

[ 'quieren', mi].

'hacemos'.

[ 'hacemos', mi].

'deteneron'.

[ 'deteneron', mi].

'hacido'.

[ 'hacido', mi].

'dirme'.

[ 'dirme', mi].

'quierian'.

[ 'quierian', mi].

'demonstran'.

[ 'demonstran', mi].

'moleron'.

[ 'moleron', mi].

'conta'.

[ 'conta', mi].

'tenimos'.

[ 'tenimos', mi].

Se observan las siguientes formas no estándares marcadas por el analizador como “formas desconocidas”( mi: mot inconnue): los verbos “quierian” (querían), “quieren” (quieren), “tenio” (tuvo), “tenimos” (tuvimos), “conta” (cuenta), “moleron” (molieron), “demonstran” (demuestran), dirme” (darme o decirme), “Hacemos” (hicimos) “quiereamos” (queremos), “refleía”(reflejaba), “tenges”(tenés), teniendo en cuenta estas palabras no analizadas y las otras observadas, se propone una implementación particular de SMORPH para el análisis de textos escritos producidos por aprendientes de español L2.

Para que reconozca las formas propias de interlengua se emplean los siguientes procedimientos:

### 10.1.1. Modificaciones en el archivo “entradas”

Para el caso de DESV2, las formas verbales conjugadas como irregulares, siendo regulares en la lengua estándar, crearemos un nuevo verbo, modificando el archivo entradas, el de rasgos y el de modelos.

Página | 142

Por ejemplo, tomamos “quiere” y “quiere” y creamos el verbo desviado *quierer* al que le asignamos un nuevo modelo, el número 63 ya que existen 62 modelos verbales.

quierer @v63 .

quier @v63 .

El mismo procedimiento puede utilizarse para otros verbos de la segunda conjugación como el caso de “podemos” (podemos):

Pueder @v63

Pued @v63

En cambio, para la DESV 3, el caso de las formas verbales con terminaciones inexistentes en el español estándar, al no ser productivas no conviene establecer un modelo de conjugación y sólo se las puede declarar en las entradas con el rasgo dsverb.

Por ejemplo, en el archivo entradas se deberán ingresar todas las formas desviadas indicando que son verbos desviados mediante el rasgo dsverb, de la siguiente manera:

demonstran /dsverb .

moleron /dsverb .

Esto implica ingresar todos los términos desviados y por lo tanto es una forma poco económica de realizar el reconocimiento.

### 10.1.2. Modificaciones en el archivo “rasgos”

Deberemos crear la etiqueta morfosintáctica que corresponda a la forma verbal desviada: dsverb, que representará a desve, así como v representa a verbo.

Página | 143

EMS	etmorsin
v	verbo
<b>dsverb</b>	<b>desve</b>
aux	auxil
ser	ser
adv	adver
nom	nombr
<b>dsvnom</b>	<b>desvi</b>
npr	npm_propio
det	determ

### 10.1.3. Modificaciones en el archivo “modelos”

Para el caso de b), creamos un nuevo modelo. Como la raíz es el lema menos los dos caracteres de la terminación del infinitivo, lo indicamos mediante -2 y luego anotamos las terminaciones correspondientes a la segunda conjugación (infinitivo en -er), junto con los siguientes rasgos: tiempo, modo, persona, número, verbo irregular y forma desviada, de la siguiente manera:

```
@v63      -2

+emos
v/pres/ind/1a/pl/c3/irr/dsverb

+éis
v/pres/ind/2a/pl/c3/irr/dsverb

+ían
v/imp/ind/3a/pl/c3/irr/dsverb
```

Para el caso de a), no agregaremos un nuevo modelo si no que modificaremos el ya existente. Si tomamos por ejemplo la forma desviada “conta” (cuenta), en el archivo “entradas” tenemos los modelos que corresponden a la raíz regular y a la raíz irregular:

contar cuent @v4 .

contar @v5 .

Nos interesa trabajar en el modelo de conjugación regular ya que la forma verbal desviada ha seguido dicha raíz. Por lo tanto, en el archivo “modelos”, agregamos al @v5, la forma verbal desviada: +a v/pres/ind/3a/sg/c1/r/dsverb

@v5 -2

```
+ar      v/infin/c1/r
+ando    v/ger/c1/r
+ado      v/part/masc/sg/c1/r
+ada      v/part/fem/sg/c1/r
+ados     v/part/masc/pl/c1/r
+adas     v/part/fem/pl/c1/r
+ás       v/pres/ind/2a/sg/c1/r/rp
+és       v/pres/subj/2a/sg/c1/r/rp
```



+amos	v/pres/ind/1a/pl/c1/r/estrpi
+áis	v/pres/ind/2a/pl/c1/r/est
+aba	v/imp/ind/1a/sg/c1/r/estrpi
+abas	v/imp/ind/2a/sg/c1/r/estrpi
+aba	v/imp/ind/3a/sg/c1/r/estrpi
+ábamos	v/imp/ind/1a/pl/c1/r/estrpi
+abais	v/imp/ind/2a/pl/c1/r/est
+aban	v/imp/ind/2a/pl/c1/r/estrpi
+aban	v/imp/ind/3a/pl/c1/r/estrpi
+é	v/prets/ind/1a/sg/c1/r/estrpi
+aste	v/prets/ind/2a/sg/c1/r/estrpi
+ó	v/prets/ind/3a/sg/c1/r/estrpi
+amos	v/prets/ind/1a/pl/c1/r/estrpi
+asteis	v/prets/ind/2a/pl/c1/r/est
+aron	v/prets/ind/3a/pl/c1/r/estrpi
+aré	v/fut/ind/1a/sg/c1/r/estrpi
+arás	v/fut/ind/2a/sg/c1/r/estrpi
+ará	v/fut/ind/3a/sg/c1/r/estrpi
+aremos	v/fut/ind/1a/pl/c1/r/estrpi
+aréis	v/fut/ind/2a/pl/c1/r/est
+arán	v/fut/ind/3a/pl/c1/r/estrpi
+aría	v/cond/ind/1a/sg/c1/r/estrpi
+arías	v/cond/ind/2a/sg/c1/r/estrpi
+aría	v/cond/ind/3a/sg/c1/r/estrpi
+aríamos	v/cond/ind/1a/pl/c1/r/estrpi
+aríais	v/cond/ind/2a/pl/c1/r/est
+arían	v/cond/ind/3a/pl/c1/r/estrpi
+ara	v/pret/subj/1a/sg/c1/r/estrpi
+emos	v/pres/subj/1a/pl/c1/r/estrpi
+éis	v/pres/subj/2a/pl/c1/r/est
+ase	v/pret/subj/1a/sg/c1/r/estrpi
+aras	v/pret/subj/2a/sg/c1/r/estrpi
+ases	v/pret/subj/2a/sg/c1/r/estrpi
+ara	v/pret/subj/3a/sg/c1/r/estrpi
+ase	v/pret/subj/3a/sg/c1/r/estrpi
+áramos	v/pret/subj/1a/pl/c1/r/estrpi
+ásemos	v/pret/subj/1a/pl/c1/r/estrpi
+arais	v/pret/subj/2a/pl/c1/r/est
+aseis	v/pret/subj/2a/pl/c1/r/est
+asen	v/pret/subj/2a/pl/c1/r/estrpi
+aran	v/pret/subj/3a/pl/c1/r/estrpi
+asen	v/pret/subj/3a/pl/c1/r/estrpi
+are	v/fut/subj/1a/sg/c1/r/estrpi
+ares	v/fut/subj/2a/sg/c1/r/estrpi
+are	v/fut/subj/3a/sg/c1/r/estrpi
+áremos	v/fut/subj/1a/pl/c1/r/estrpi

+areis v/fut/subj/2a/pl/c1/r/est  
 +aren v/fut/subj/3a/pl/c1/r/estrpi  
 +á v/pres/imper/2a/sg/c1/r/rp  
 +émos v/pres/imper/1a/pl/c1/r/estrpi  
 +emos v/pres/imper/1a/pl/c1/r/estrpi  
 +ad v/pres/imper/2a/pl/c1/r/est .

**+a v/pres/ind/3a/sg/c1/r/dsverb**

De este modo, la herramienta analizará como 3era persona singular tanto a *cuenta* como a *conta*, distinguiendo a ésta última como forma desviada.

Si tomamos “tenimos” (tuvimos), “tenio”, (tuvo), realizamos la misma operación: en el modelo de la raíz regular: tener @v42 agregamos las formas verbales desviadas:

**+io**  
**v/prets/ind/3a/sg/dsverb**  
  
**+imos**  
**v/prets/ind/1a/pl/dsverb**

@v42 -2

+er v/infin/estrpi  
 +iendo v/ger/estrpi  
 +ido v/part/masc/sg/estrpi  
 +ida v/part/fem/sg/estrpi  
 +idos v/part/masc/pl/estrpi  
 +idas v/part/fem/pl/estrpi  
 +és v/pres/ind/2a/sg/rp  
 +emos v/pres/ind/1a/pl/estrpi  
 +éis v/pres/ind/2a/pl/est  
 +ía v/imp/ind/1a/sg/estrpi  
 +ías v/imp/ind/2a/sg/estrpi  
 +ía v/imp/ind/3a/sg/estrpi  
 +íamos v/imp/ind/1a/pl/estrpi  
 +íais v/imp/ind/2a/pl/est  
 +ían v/imp/ind/3a/pl/estrpi

+é      v/pres/imper/2a/sg/rp  
+ed      v/pres/imper/2a/pl/est .

**+imos      v/prets/ind/1a/pl/dsverb**

**+io      v/prets/ind/3a/sg/dsverb**

#### 10.1.4 Resultados del Análisis de Smorph

Smorph presenta la etiqueta morfosintáctica ('EMS') verbo ('v'), aclara que es una forma verbal desviada ('dsverb'), luego que se trata del modo indicativo ('MODOV', 'ind'), nos informa sobre la persona (1ª, 2ª, 3ª) y el número (singular "sg" o plural "pl") ('PERS', '1ª'), ('NUM', 'pl'), a continuación el tiempo (presente "pres", imperfecto "imp") (TPO 'pres') y por último aclara si el verbo es regular ('TR', 'r') o irregular ('TIRR') y en este caso brinda rasgos de la irregularidad.

'conta'.

[ 'contar', 'EMS','v', 'EMS','**dsverb**', 'MODOV','ind', 'PERS','3a', 'NUM','sg', 'TPO','pres', 'TR','r', 'TC','c1'].

'tenio'.

[ 'tener', 'EMS','v', '**EMS**','**dsverb**', 'MODOV','ind', 'PERS','3a', 'NUM','sg', 'TPO','prets'].

'tenimos'.

[ 'tener', 'EMS','v', '**EMS**','**dsverb**', 'MODOV','ind', 'PERS','1a', 'NUM','pl', 'TPO','prets'].

'quiereamos'.

[ 'quiereamos', 'EMS','v', '**EMS'**, '**dsverb'**, 'MODOV','ind', 'PERS','1a', 'NUM','pl', 'TPO','pres', 'TR','irr', 'TC','c3', 'TDIAL',].

[ 'quierer', 'EMS','v', '**EMS'**, '**dsverb'**, 'MODOV','ind', 'PERS','1a', 'NUM','pl', 'TPO','pres', 'TR','irr', 'TC','c3',].

'quierían'.

[ 'quierían', 'EMS','v', '**EMS'**, '**dsverb'**, 'MODOV','ind', 'PERS','3a', 'NUM','pl', 'TPO','imp', 'TR','irr', 'TC','c3',].

[ 'quierer', 'EMS','v', '**EMS'**, '**dsverb'**, 'MODOV','ind', 'PERS','3a', 'NUM','pl', 'TPO','imp', 'TR','irr', 'TC','c3',].

'demonstran'.

[ 'demonstran', '**EMS'**, '**dsverb'**].

'moleron'.

[ 'moleron', '**EMS'**, '**dsverb'**].

Para el caso de DESVIACIÓN 4, formas verbales que se construyen con terminaciones existentes en la lengua estándar pero que no corresponden a esas bases verbales, podemos reconocerlas a partir de las terminaciones verbales, especificando mediante las terminaciones distinguidas, que

toda palabra que tenga una terminación verbal sea considerada verbo.

La dificultad estará en que no se distinguirá si son formas verbales de la lengua estándar o de la interlengua, es decir, no podemos advertir que son formas verbales desviadas.

## 10.2. Implementación en Nooj

El sistema Nooj tiene la ventaja de efectuar el análisis de textos a partir de la indicación acerca de los diccionarios y las gramáticas asociadas a éstos con los que debe trabajar, por eso es factible crear gramáticas que actúen sólo para los casos de formas desviadas de la interlengua y de esta forma no modificar la información lingüística declarada en los diccionarios y gramáticas pertenecientes al módulo español.

### 10.2.1. Generación de gramáticas productivas

Para el caso de formas verbales irregulares que son conjugadas por los aprendientes como regulares. Por ejemplo en los verbos que diptongan: “queren” (quieren), “encontra” (encuentra), “empeza” (empieza) se genera una gramática productiva que reconozca que las raíces regulares asociadas con un conjunto cualquiera de letras es un verbo de la interlengua.

Esto se indica mediante la etiqueta V (verbo)+ INT (interlengua) y además se agrega + DIPT para señalar las características de las formas estándar.

Para ello se realiza la gramática derivacional mediante grafos de la siguiente manera: se colocan las raíces regulares del modelo de verbos correspondiente (en este caso quer\_ para querer, pod\_ para poder y acert para acertar, éste último representa a los verbos de la primera conjugación), y se indica que sumado a cualquier conjunto de caracteres dé como resultado V+ INT+ DIPT.

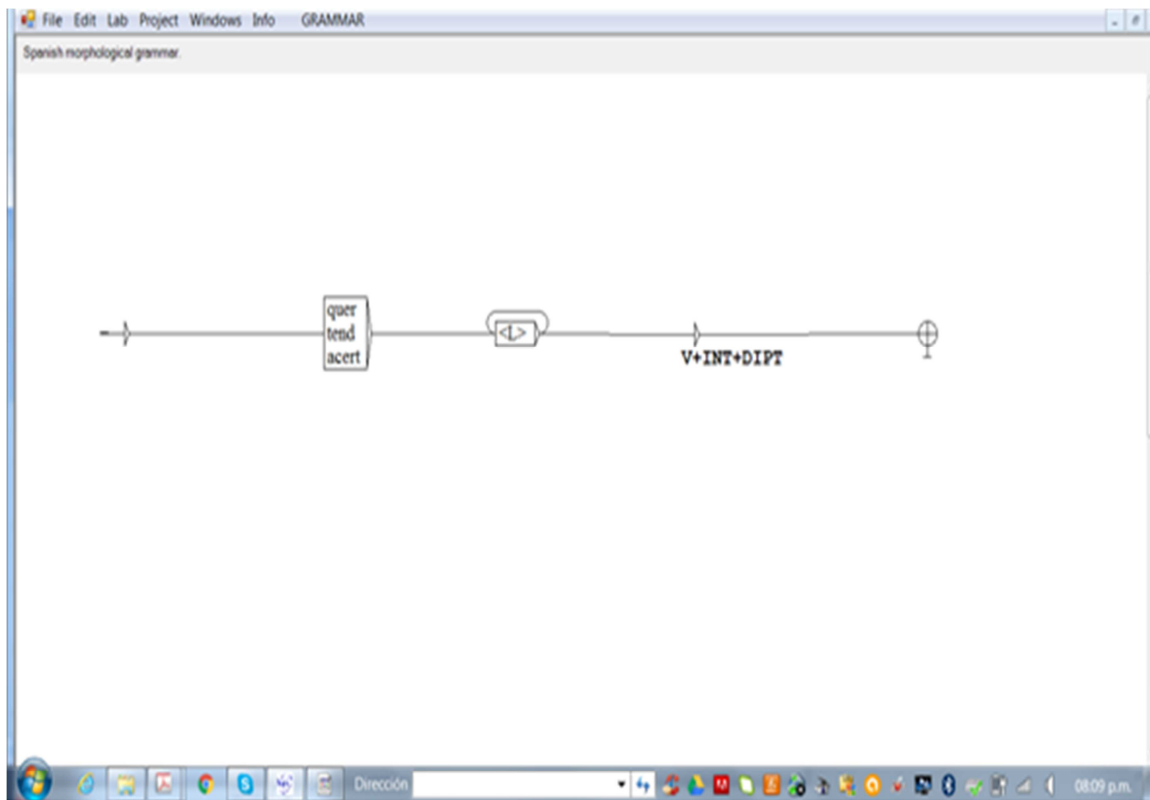


Imagen Captura de pantalla N° 8 : Gramática productiva para la desviación 1.

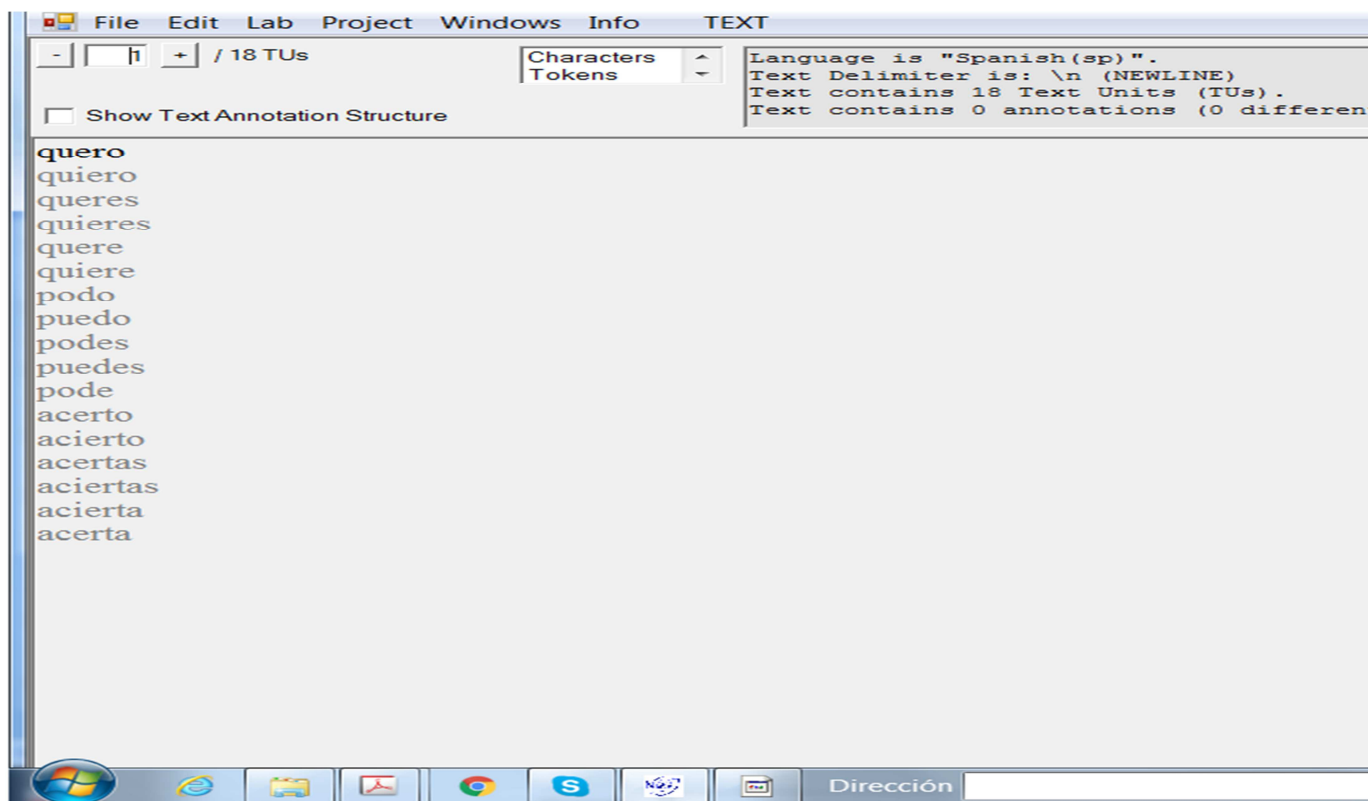
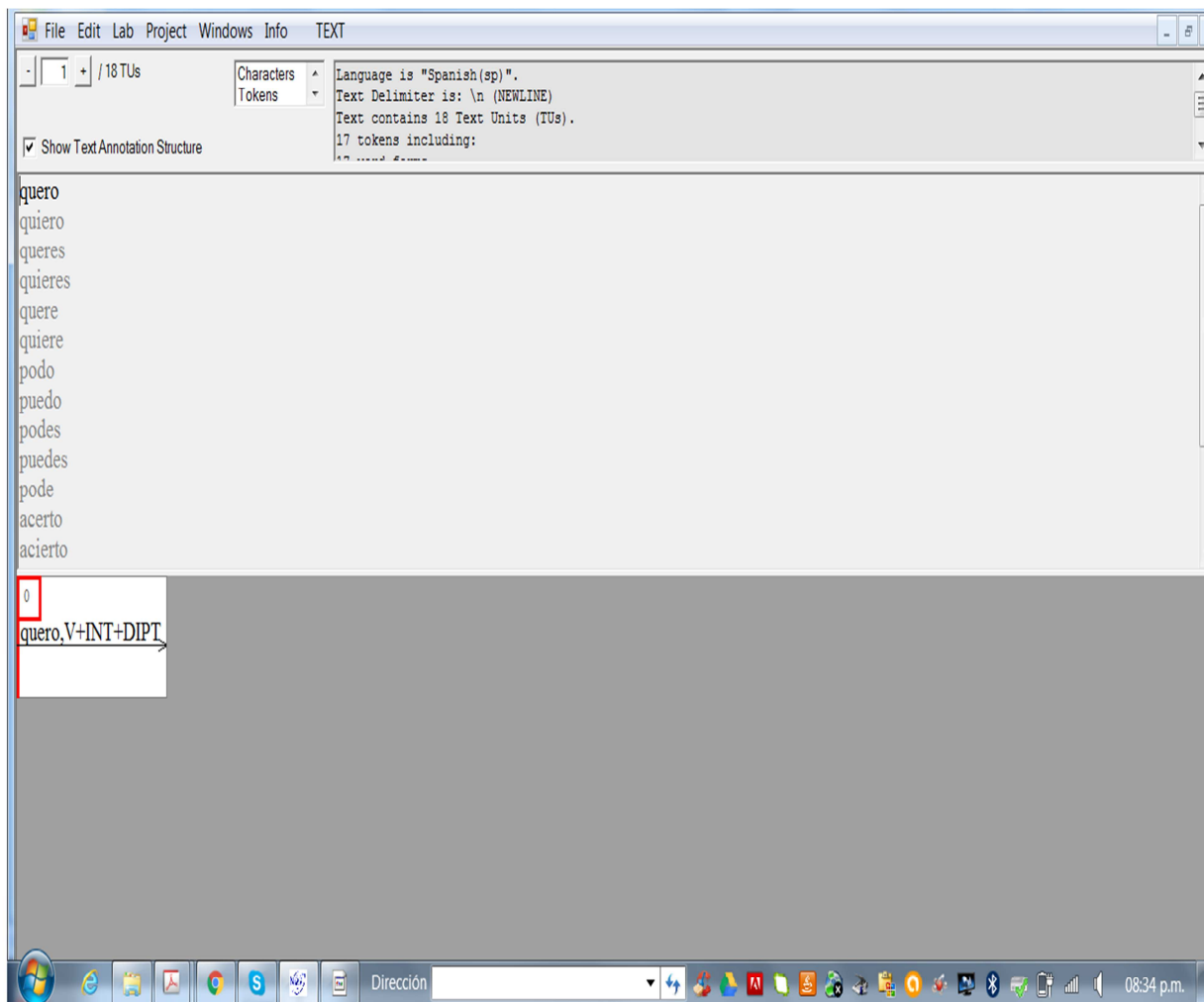


Imagen captura de pantalla N°9: Texto de prueba.

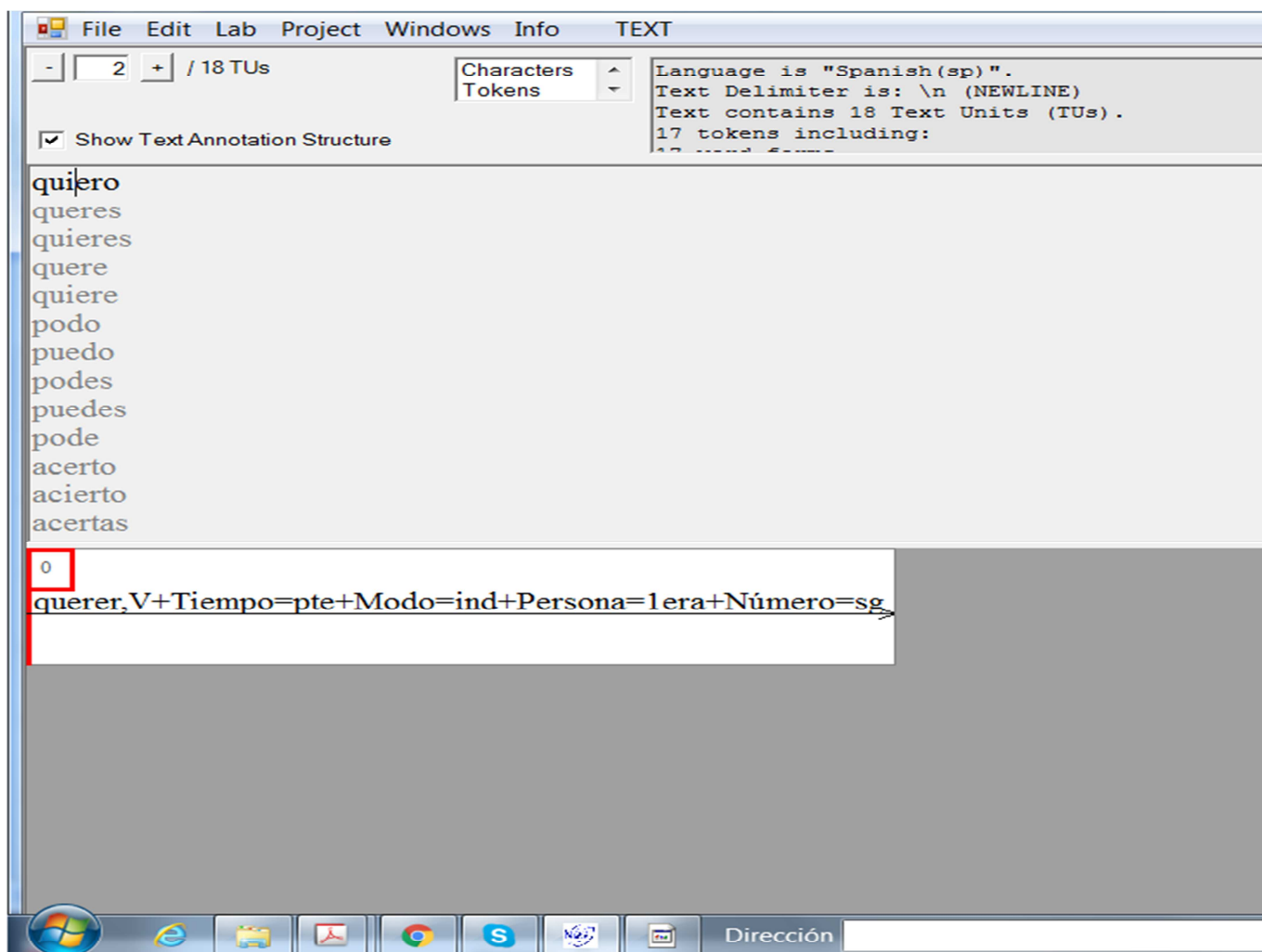
Es importante aclarar que esta gramática se aplica junto con los diccionarios del español estándar y se indica que tiene "Priority" L1 en la pestaña Lexical Analysis, cuando se seleccionan los diccionarios y gramáticas con los que debe analizar el texto. Si la palabra está en el diccionario estándar, la reconoce; si no está, la etiqueta como V+INT.



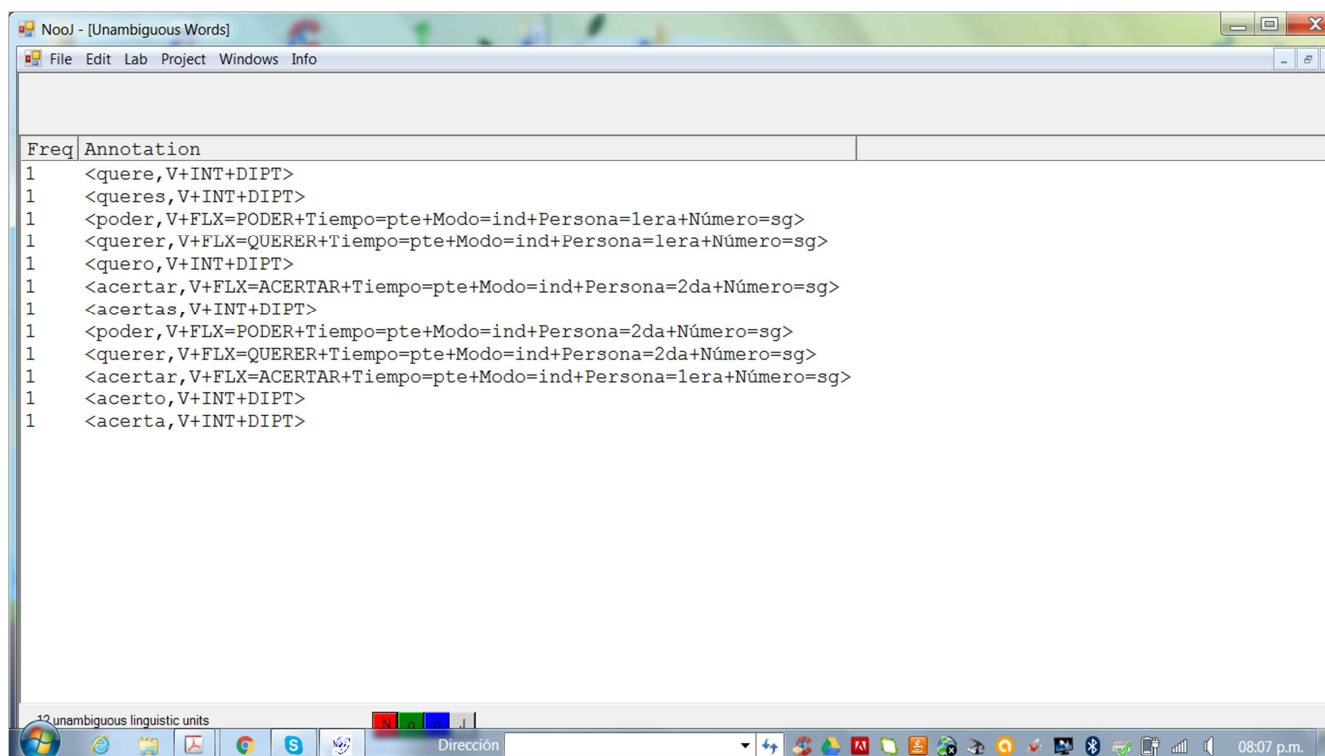
na | 152

**Captura de pantalla Nº 10: Muestra de análisis de forma verbal de interlengua.**





Captura de pantalla Nº 11: Muestra de análisis de forma verbal estándar.



Freq	Annotation
1	<quere, V+INT+DIPT>
1	<quieres, V+INT+DIPT>
1	<poder, V+FLX=PODER+Tiempo=pte+Modo=ind+Persona=1era+Número=sg>
1	<querer, V+FLX=QUERER+Tiempo=pte+Modo=ind+Persona=1era+Número=sg>
1	<quero, V+INT+DIPT>
1	<acertar, V+FLX=ACERTAR+Tiempo=pte+Modo=ind+Persona=2da+Número=sg>
1	<acertas, V+INT+DIPT>
1	<poder, V+FLX=PODER+Tiempo=pte+Modo=ind+Persona=2da+Número=sg>
1	<querer, V+FLX=QUERER+Tiempo=pte+Modo=ind+Persona=2da+Número=sg>
1	<acertar, V+FLX=ACERTAR+Tiempo=pte+Modo=ind+Persona=1era+Número=sg>
1	<acerto, V+INT+DIPT>
1	<acerta, V+INT+DIPT>

12 unambiguous linguistic units

#### Captura de pantalla Nº 12 : Análisis formas no ambiguas

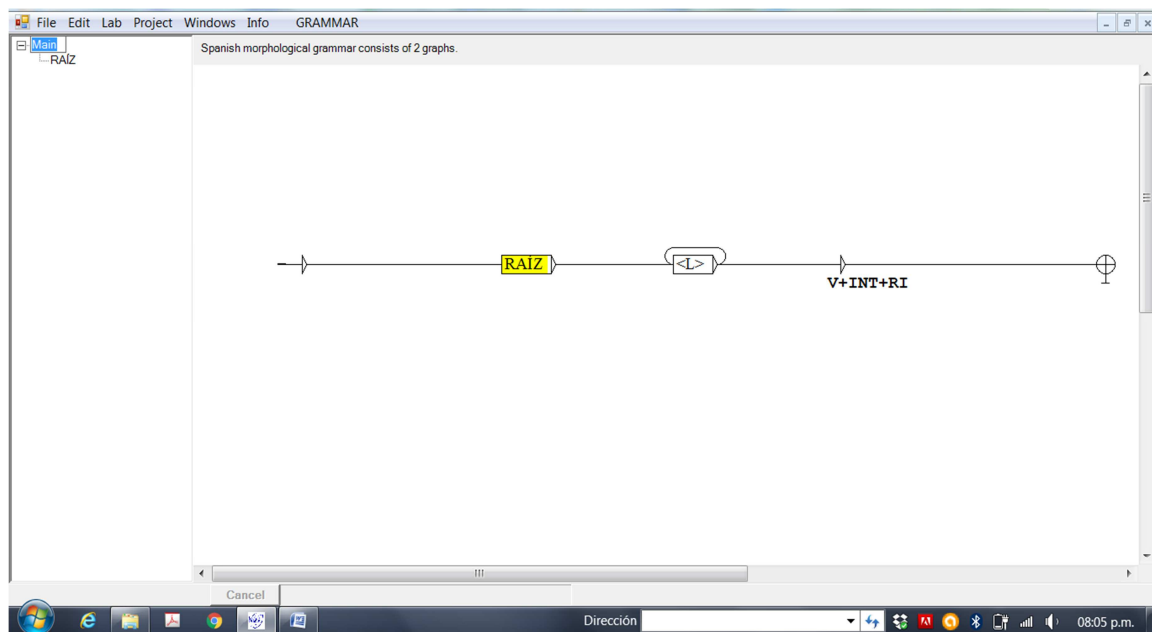
Ahora bien, mediante esta gramática productiva también podría reconocer el otro grupo de casos dentro de las desviaciones del tipo 1 que son los verbos que presentan cambios gráficos, colocando que las raíces que no presentan variaciones (como por ejemplo sac\_ localiz- busc-) asociadas a cualquier desinencia verbal, dé como resultado un verbo de interlengua. El problema aquí radica en que ya no podríamos utilizar la etiqueta+ DIPT.

Lo más conveniente y que abarcaría a una mayor proporción de formas verbales irregulares desviadas sería crear una etiqueta como "V+INT+RI", (verbo de interlengua que debe llevar raíces irregulares) con lo cual estaríamos indicándole al aprendiente que tiene que emplear la raíz irregular.

De este modo analizaría casos como sugirió (sugirió), proponí (propuse) sacé (saqué) pedo (pido).

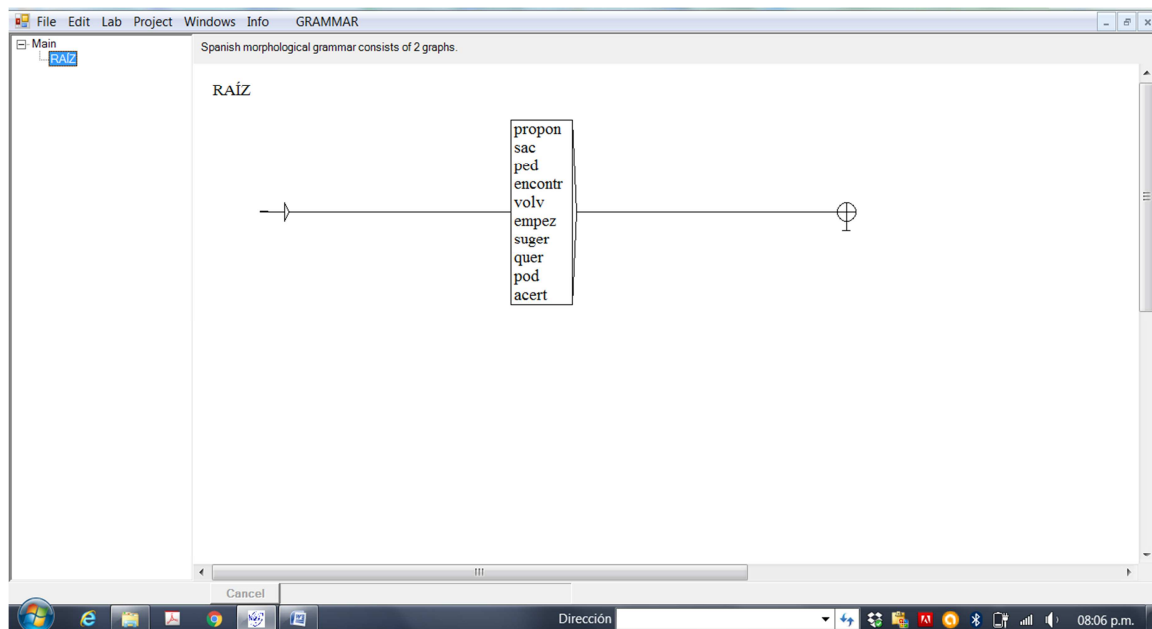
Para generar esta gramática debemos indicar que está compuesta por dos grafos: en el primero consignamos las raíces regulares y en el segundo cualquier secuencia de caracteres para que dé como resultado la etiqueta "V+INT+RI".

En las imágenes que siguen observamos la gramática productiva mediante grafos:

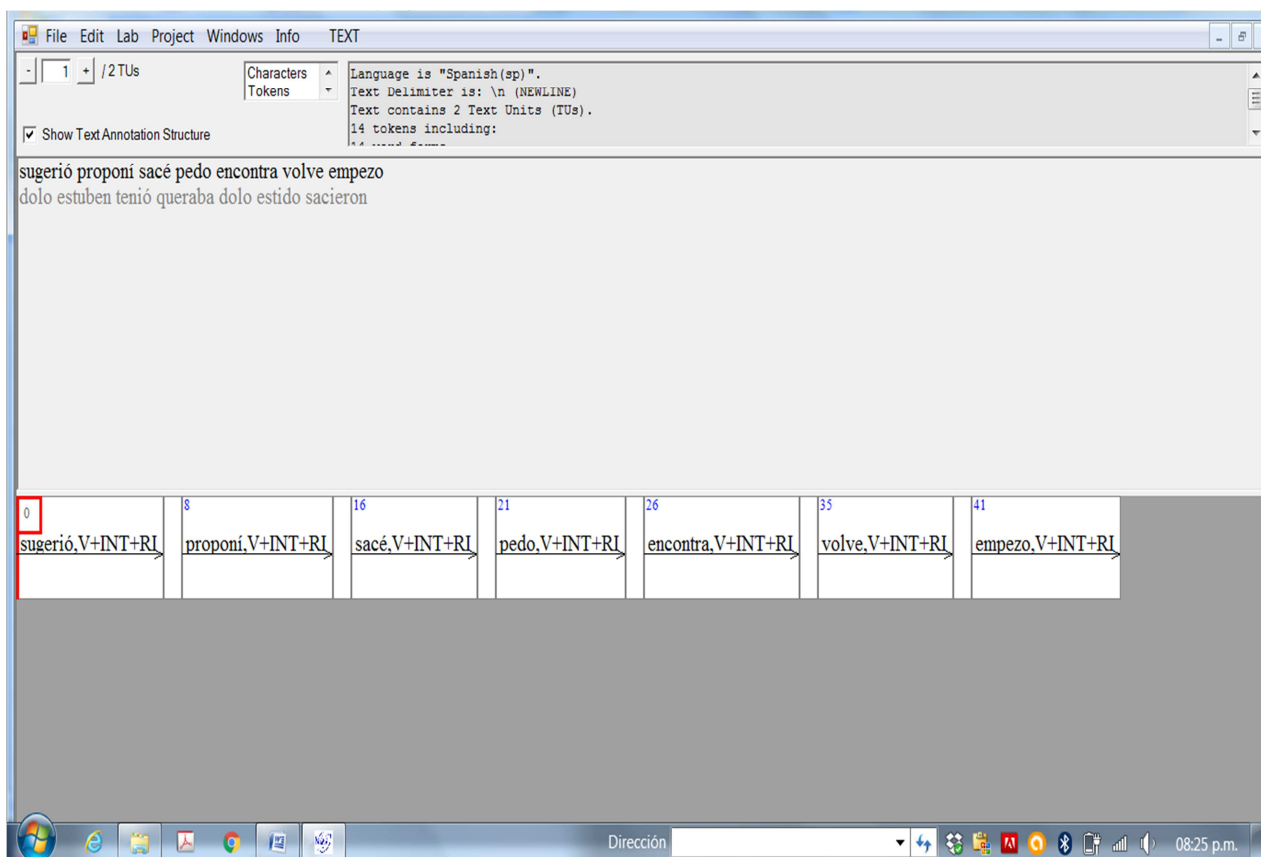


Página | 155

Captura de pantalla N° 13: gramática productiva para desviación tipo 1



Captura de pantalla N°14: gramática productiva para desviación 1(raíces regulares)

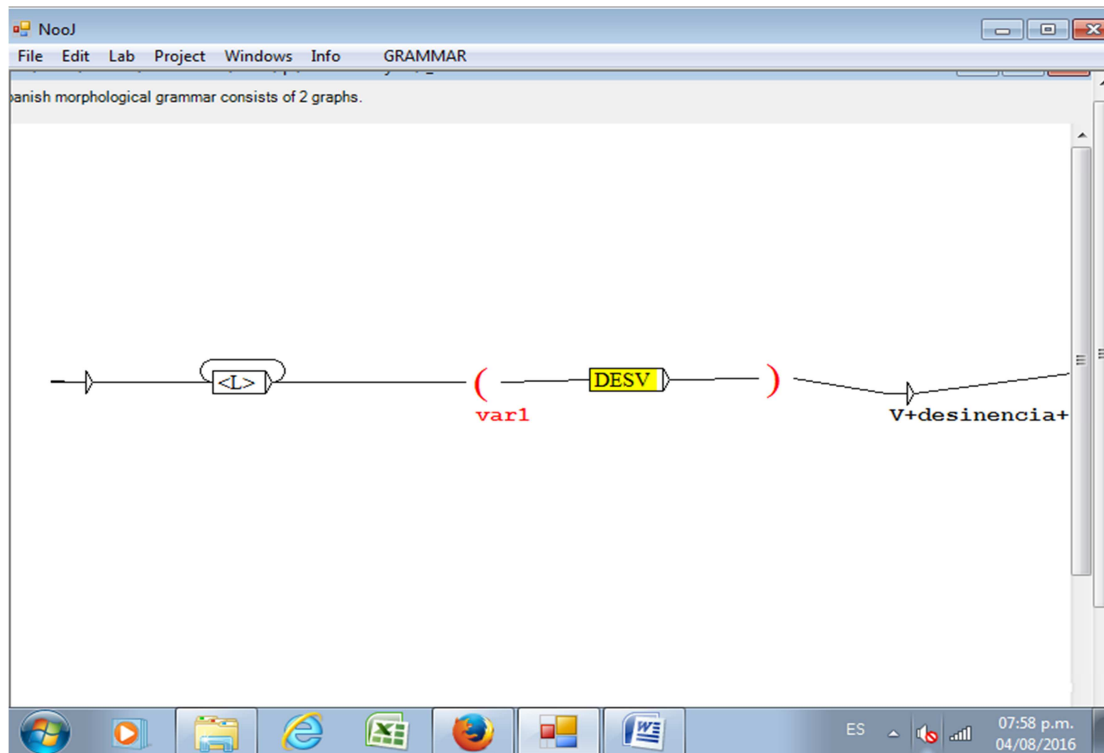


156

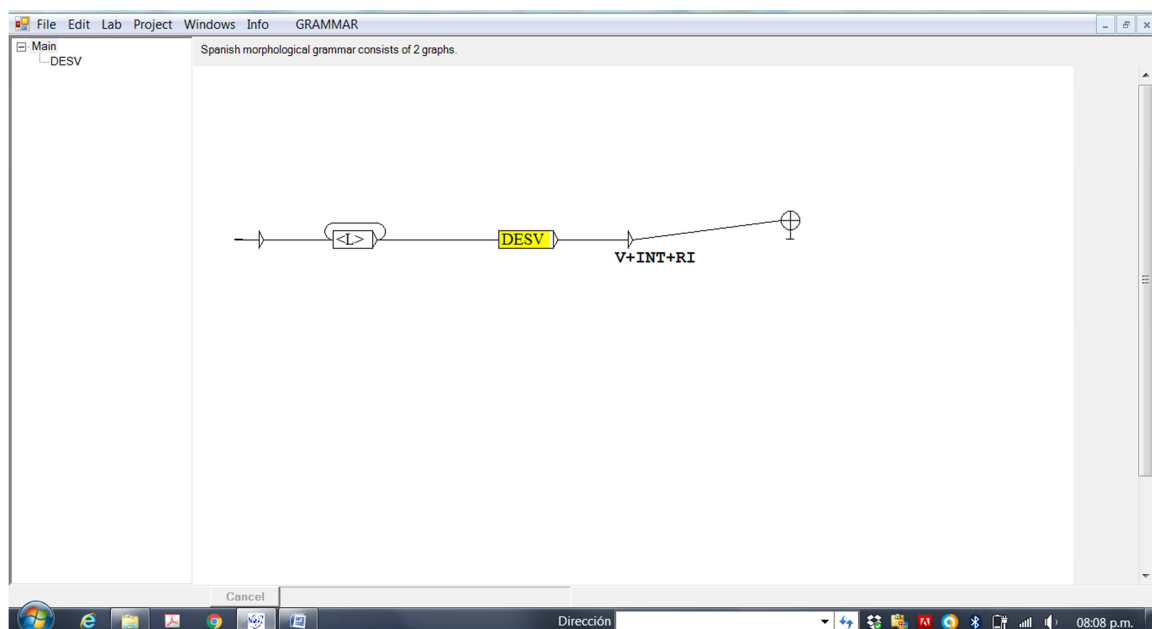
### Captura de pantalla N°15: resultado análisis formas verbales desviadas 1

Para el caso de desviaciones del tipo 4, formas verbales que presentan una asignación incorrecta de desinencia, es decir, se construyen con terminaciones existentes en la lengua estándar pero que no corresponden a esas bases verbales, podría generarse otra gramática productiva que indique que cualquier secuencia de caracteres asociada con desinencias verbales (en este caso podrían colocarse las de presente y las de pretérito del Modo Indicativo, por ser los tiempos verbales que predominan en el corpus de interlengua) es un verbo de interlengua.

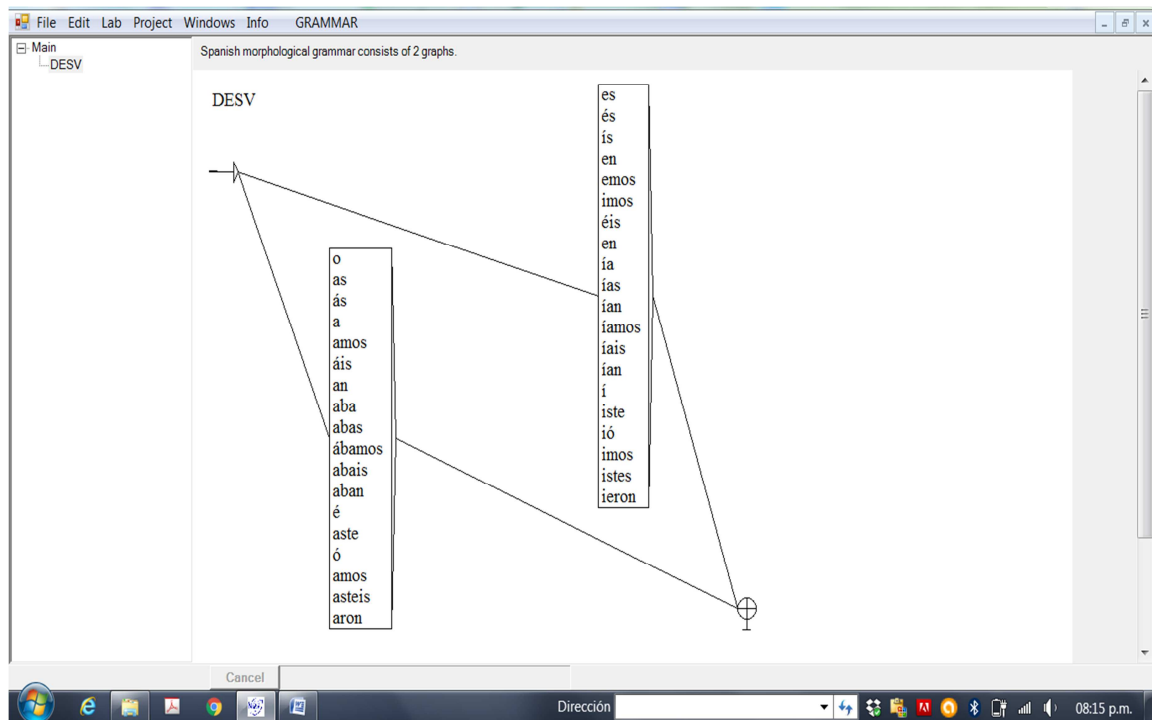
Las siguientes imágenes muestran la gramática generada mediante grafos:



Captura de pantalla N°16: gramática productiva para desv 4 (grafos) (2)



Captura de pantalla N° 17: gramática productiva para desv 4 (grafos) (1)



Captura de pantalla Nº 18: gramática productiva para desv4 (desinencias verbales)

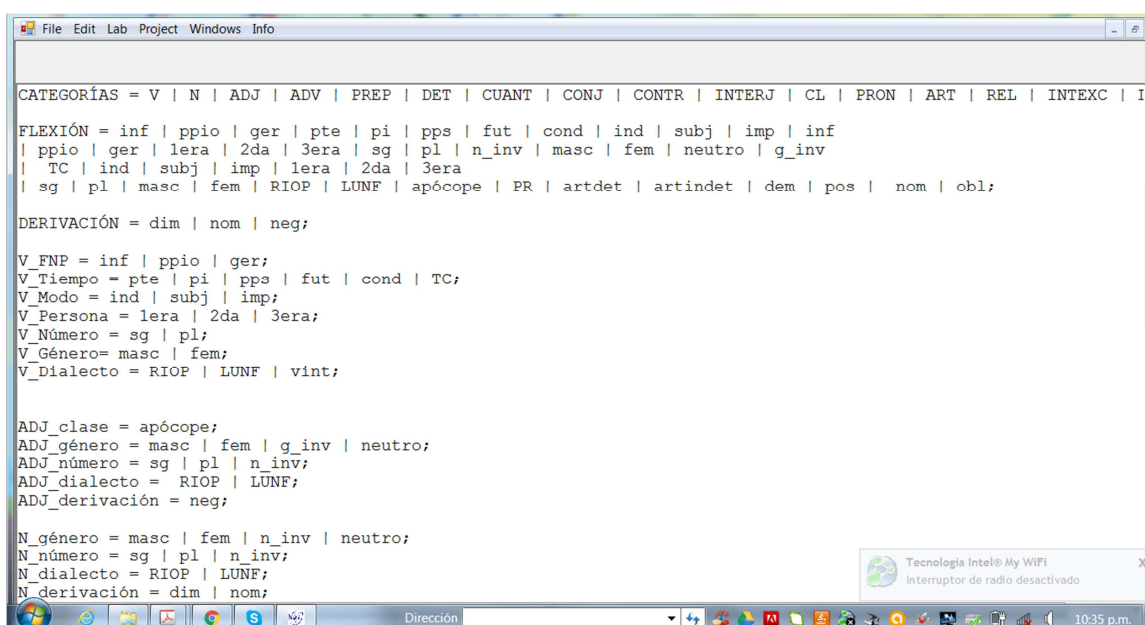
Token	Verb Form	Grammatical Category
0	dolo	V+INT+RI
5	estuben	V+INT+RI
13	tenió	V+INT+RI
19	queraba	V+INT+RI
27	dolo	V+INT+RI
32	estido	V+INT+RI
39	sacieron	V+INT+RI

Captura de pantalla Nº19: resultado análisis formas verbales desviadas 4

El problema radicaría en la precisión del análisis debido a que las desinencias son muy inespecíficas (o, es, e, emos, éis, en) y va a reconocer como verbo de interlengua cualquier palabra que no esté en el diccionario y termine con esos caracteres.

En cuanto a las desviaciones del tipo 2, cuando las formas verbales regulares son conjugadas por los aprendientes como irregulares generaríamos una gramática productiva que diga que las raíces irregulares asociadas a un conjunto cualquiera de letras es un verbo de interlengua mediante la siguiente etiqueta: "V+INT+RR", con lo cual estaríamos indicándole al aprendiente que debe emplear la raíz regular para que su verbo coincida con el de la lengua estándar. De esta manera detectaría y etiquetaría formas como: queremos (queremos), podemos (podemos) rueguía (rogaba).

Para las desviaciones del tipo 3, las que escapan a toda sistematicidad por presentar intercambio de vocales o consonantes o repetición de grafemas, por ejemplo: *tien*, *he empezato*, *hacemos*, *arrepintinron*, *estarei*, la única alternativa es crear un diccionario propio de verbos de interlengua e ingresarlas una por una, para eso previamente habrá que crear una categoría como por ejemplo "vint" en el archivo "\_properties.def", y agregar este valor en "V\_dialecto".



```

File Edit Lab Project Windows Info

CATEGORÍAS = V | N | ADJ | ADV | PREP | DET | CUANT | CONJ | CONTR | INTERJ | CL | PRON | ART | REL | INTEXC | I

FLEXIÓN = inf | ppio | ger | pte | pi | pps | fut | cond | ind | subj | imp | inf
| ppio | ger | lera | 2da | 3era | sg | pl | n_inv | masc | fem | neutro | g_inv
| TC | ind | subj | imp | lera | 2da | 3era
| sg | pl | masc | fem | RIOP | LUNF | apócope | PR | artdet | artindet | dem | pos | nom | obl;

DERIVACIÓN = dim | nom | neg;

V_FNP = inf | ppio | ger;
V_Tiempo = pte | pi | pps | fut | cond | TC;
V_Modo = ind | subj | imp;
V_Persona = lera | 2da | 3era;
V_Número = sg | pl;
V_Género = masc | fem;
V_Dialecto = RIOP | LUNF | vint;

ADJ_clase = apócope;
ADJ_género = masc | fem | g_inv | neutro;
ADJ_número = sg | pl | n_inv;
ADJ_dialecto = RIOP | LUNF;
ADJ_derivación = neg;

N_género = masc | fem | n_inv | neutro;
N_número = sg | pl | n_inv;
N_dialecto = RIOP | LUNF;
N_derivación = dim | nom;

Tecnología Intel® My WiFi
Interruptor de radio desactivado

Dirección 10:35 p.m.

```

**Captura de pantalla N° 20: valor vint en el archivo properties.def.**

El diccionario para verbos de interlengua se consigna de la siguiente manera:

tien, V_dialecto vint
embarquei, V_dialecto vint
hacemos, V_dialecto vint
arrepintinron, V_dialecto vint

Sin embargo, gran parte de las formas desviadas agrupadas bajo la categoría desv3 pueden ser analizadas con la gramática generativa que generé para los casos de desv4, indicando que etiquete como V+INT a cualquier serie de caracteres unidas a desinencias verbales.

Reconoce formas como: *piqe, agregé, coloque, concí, ha, demonstnan, trayera, trajese, pensio, dares, agaría, pusque, enquentre, cuosta, estive, demostré, tibe, pondiera*.

De esta forma cubriría el 65 %de los casos de desv3.

### 10. 3. Conclusiones

Luego de realizar las adaptaciones en los programas Smorph y Nooj con el objetivo de detectar las formas verbales irregulares desviadas, puedo concluir que el sistema Nooj cuenta con la ventaja de trabajar con diccionarios y gramáticas que pueden funcionar en conjunto o individualmente para analizar los textos.

En otras palabras, el sistema permite la creación de diccionarios y gramáticas (tanto derivacionales como sintácticas) exclusivas de interlengua.

Esto es muy conveniente a la hora de analizar textos de aprendientes de español ya que no implica modificar la información lingüística ya declarada en los archivos, como ocurre con el generador y analizador Smorph.



No obstante, sí requiere trabajar en conjunto con los diccionarios de español, precisamente para distinguir las formas pertenecientes a la lengua estándar de las que son idiosincrásicas. Cabe aclarar que la creación de modelos verbales irregulares que está llevando a cabo mi equipo de trabajo para el módulo de español de Nooj fue posible gracias a la información lingüística declarada en el software Smorph, sobre todo en lo que atañe a la clasificación de verbos irregulares y a su factorización según puedan agruparse de acuerdo a las terminaciones verbales, como mostré en el capítulo 9.

Para los casos de desv1 en Smorph tuve que incorporar las desinencias correspondientes en cada modelo verbal y así modificar el modelo de querer, tener, poder, pensar, entre otros. En cambio, en Nooj creando sólo una gramática productiva que tome las raíces regulares de los verbos más las desinencias (ambos ya están implantadas en los archivos correspondientes a la morfología verbal del español), conseguí detectar la totalidad de las formas desviadas del tipo 1, abarcando tanto a verbos que diptongan, como a los que presentan cambios vocálicos por cierre de vocal y a los que llevan cambios gráficos.

Cabe aclarar que no sólo logró analizarlas automáticamente sino que además pudo proporcionar una etiqueta (RI) con el fin de indicar que debe emplearse la raíz irregular de ese verbo para formar un verbo que pertenezca al español estándar.

Con respecto al grupo de desv2, aunque éstas sean muy escasas (recordemos que sólo representan un 3% del total de formas verbales irregulares) pude reconocerlas en Nooj mediante la creación de una gramática productiva en la que se consignan todas las raíces irregulares y se aclara que seguidas de cualquier secuencia de caracteres dé como etiqueta "V+INT+RR". Este último rasgo (RR) sirve para señalar que deben emplearse las raíces regulares de esos verbos en lugar de las irregulares.

Por lo tanto, para los primeros dos tipos de desviaciones, fue exitosa la medición en términos de cobertura y precisión<sup>23</sup>, alcanzando los valores esperados en el 100% de los casos.

En Smorph también fue posible el reconocimiento automático de estos dos grupos de desviaciones (1 y 2) pero alterando los modelos verbales de los verbos irregulares para agregar las formas desviadas para desv1 y además creando verbos inexistentes en español, propios de interlengua en el archivo entradas como por ejemplo *quierer* para las desv2.

En cuanto al grupo de desv4, en Smorph no obtuve resultados positivos en términos de precisión ya que de emplear las desinencias como terminaciones distinguidas para reconocer estos casos, el programa detectaría como verbos a toda palabra que terminara con dichas desinencias pero no discriminaría si se trata de verbos pertenecientes al español estándar o verbos de interlengua, provocando así gran cantidad de errores.

En Nooj, logré detectarlos gracias a la creación de una gramática morfológica que reconoce como verbo de interlengua a toda secuencia de caracteres que termine con desinencias verbales.

Los resultados en cuanto a cobertura son del 100%. ¿Pero qué sucede con respecto a la precisión? ¿Se corre el riesgo de que reconozca como verbo de interlengua a cualquier palabra cuya terminación coincida con una desinencia verbal?

Para que no etiquete a palabras cuya terminación coincide con una desinencia verbal como por ejemplo helado que termina con o y coincide con la terminación de primera persona singular del presente Modo indicativo, se deben seleccionar además del diccionario de verbos de español, los diccionarios que involucran a todas las demás categorías, es decir, el de nombres, adjetivos, adverbios, preposiciones, pronombres, entre otros.

---

<sup>23</sup> La precisión se define como el cociente entre el número de ítems correctamente analizados y el número total de ítems analizados, mientras que la cobertura es el cociente entre el número de ítems correctamente analizados y el número total de ítems.

El único problema sería que analice como "V+INT+RI" a formas verbales de interlengua que no son irregulares, en el caso de que existieran en el texto desviaciones en formas verbales regulares. Sin embargo, en el corpus analizado no hallé anomalías en la flexión de verbos regulares.

Por lo tanto, se obtienen los mismos resultados en cuanto a cobertura y precisión.

Para finalizar, nos quedan las desv3, que se tornan las más difíciles de reconocer ya que por sus características escapan a toda sistematización.

En Smorph debí ingresar todos los ítems léxicos en las entradas con el rasgo desverb. Sin embargo esta forma resultó poco económica y poco productiva, significando además una sobrecarga en las entradas. Por consiguiente arroja una buena precisión pero no permite reconocer las formas desviadas si antes no se las ingresa de la forma que expliqué.

Por lo tanto, y para poder analizar cualquier texto de interlengua fue necesario buscar una manera de abarcar la mayor cantidad posible de formas desviadas. La solución al problema la encontré mediante la utilización de la gramáticas morfológica formulada para detectar los casos de desv4.

Con la aplicación de la misma gramática y los diccionarios del módulo español no sólo puede reconocer las desv4 sino también la mayoría de las desv3.

En resumen, el sistema Nooj no sólo permitió realizar el reconocimiento automático de una manera mucho más económica, sino que además proporcionó una mayor cobertura, abarcando los casos desv4, y el 65% de desv3 a partir de la creación de una gramática morfológica que reconoce que cualquier conjunto de caracteres más desinencias verbales es un verbo de interlengua.

La siguiente tabla muestra el alcance de ambas herramientas en el reconocimiento automático de los cuatro tipos de desviaciones:

	<b>SMORPH</b>	<b>NOOJ</b>
<b>DESV1</b>	Se logra de forma poco económica y requiere alterar los modelos verbales de español.	Cobertura y precisión 100%
<b>DESV2</b>	Se logra de forma poco económica e implica la generación de verbos inexistentes en español.	Cobertura y precisión 100%
<b>DESV3</b>	Precisión sin cobertura	Cobertura 65% precisión 80%
<b>DESV4</b>	Cobertura sin precisión	Cobertura y precisión 100%

**Tabla Nº 58: Alcances de las herramientas informáticas**

En la presente investigación se llevó a cabo un estudio sobre la flexión verbal irregular del español que permitió la descripción y clasificación de formas verbales irregulares desviadas pertenecientes a la interlengua de aprendientes de español como L2, para su posterior reconocimiento automático mediante las herramientas Smorph y Nooj.

### Etapas de la investigación

Comencé por conformar un corpus de textos escritos de aprendientes de español, seleccionando sujetos que se encontraran en un nivel inicial de aprendizaje de la lengua española y cuyas de lenguas de origen fueran distintas. Por lo tanto reuní cinco corpus divididos según la lengua materna de los estudiantes: alemán, francés, holandés, inglés y portugués. Es importante aclarar que los sujetos pertenecientes al corpus portugués se encuentran en un nivel de aprendizaje inferior que el resto de los aprendientes.

Situando mi investigación dentro de los estudios de interlengua, luego de digitalizar el corpus y analizarlo, **observé un fenómeno particular: desviaciones en formas flexionadas de verbos irregulares.** Me propuse entonces analizar la morfología verbal irregular, extrayendo la totalidad de las formas verbales irregulares, es decir, atendiendo tanto a las formas idiosincrásicas como las coincidentes con el español estándar.

La metodología de trabajo entonces consistió en un primer momento en un estudio cualitativo.

En primer lugar estudié los modelos de análisis en el campo de la investigación de segundas lenguas, consulté bibliografía especializada sobre interlengua y la distinción de “errores” según su origen en interlingüísticos e intralingüísticos.

Establecí entonces una **hipótesis: las desviaciones observadas en la interlengua respecto de la flexión verbal irregular son de índole intralingüística**, es decir, refieren a una problemática del español y no tienen su causa en la lengua materna de los aprendientes.

Estudié las cuestiones referidas al sistema verbal irregular español, basándome en la descripción que realiza Alcoba (1999) y consultando la Gramática de la Real Academia Española (2009) a saber:

- Cuándo un verbo es irregular.
- Qué extensión abarca la irregularidad.
- ¿Cómo son las irregularidades?
- ¿Qué verbos presentan irregularidades especiales?

A Partir de esto **clasifiqué los verbos irregulares hallados en el corpus** de interlengua en cinco grupos teniendo en cuenta las variaciones que los afectan:

Irregularidad A: cambios consonánticos por alternancia o por incremento.

Irregularidad B: cambios en la raíz por diptongación.

Irregularidad C: cambios en la raíz por cierre de vocal.

Irregularidad D: cambios gráficos.

Irregularidad E: irregularidades especiales.

Además especifiqué los verbos irregulares utilizados en cada corpus según la lengua materna.

Luego **agrupé en cuatro casos de desviaciones a las formas verbales irregulares propias de interlengua**, (aquellas que no pertenecen ni a la lengua materna ni a la lengua meta pero que sin embargo mantienen características comunes con ambos sistemas lingüísticos.

Con el fin de clasificar las desviaciones las dividí en los siguientes grupos:

a) Formas verbales irregulares que son conjugadas por los aprendientes como regulares.

b) Formas verbales regulares que son conjugadas por los aprendientes como irregulares.

c) Formas verbales que presentan una irregularidad distinta a la estándar.

d) Formas verbales que se construyen con desinencias existentes en español pero que no corresponden a esas bases verbales.

### Evaluación de la hipótesis lingüística

La validación de la hipótesis lingüística referida a que las desviaciones en la morfología verbal irregular flexiva responden a una causa intralingüística implicó un estudio cuantitativo además del cualitativo ya realizado.

Para esto realicé tablas dinámicas en Excel para cada corpus de producciones según la lengua materna consignando la siguiente información: cantidad de textos por corpus, desviaciones y formas coincidentes con el español estándar halladas en cada texto, clasificación de esas ocurrencias según correspondieran a: desv1, desv2, desv3, desv4 o est y verbo irregular al que pertenecen dichas formas flexionadas.

Con estos datos realicé un estudio comparativo mediante la prueba estadística no paramétrica de Kruskal Wallis, tomando la decisión de no incluir en el análisis estadístico al corpus holandés por presentar una cantidad inferior de textos.

La prueba arrojó como resultado que los corpus: alemán, inglés y francés no diferían en cuanto a porcentaje de desviaciones 1, 2, 3, 4 y formas estándar.

Con respecto al corpus portugués tampoco presentó diferencias en cuanto a desviaciones del tipo 1, 2 y 4, mientras que sí exhibió una cantidad menor de formas coincidentes con la lengua estándar y una cantidad mayor de desviaciones del tipo 3 con respecto al resto de los corpus.

**Realicé además una tabla dinámica para determinar la frecuencia de los verbos irregulares** y las conclusiones fueron las siguientes: el verbo *ser* resultó ser el más utilizado, seguido por el verbo *tener*.

Con respecto a los verbos irregulares que muestran mayor cantidad de desviaciones: en primer lugar se ubicó el verbo *tener* (que sufre cambios consonánticos y vocálicos en el presente: tengo-tiene, es de pretérito fuerte: tuve-tuvieron y además varía en el futuro: tendré) y en segundo lugar, el verbo *estar* (que presenta la particularidad en presente de flexionarse como *estoy* y es de pretérito fuerte: estuve-estuvieron). Todos ellos pertenecen según la clasificación propuesta, a los verbos de irregularidades especiales.

Finalmente, se pudo comprobar que el corpus portugués es el que posee un mayor porcentaje de formas verbales desviadas de los cuatro corpus comparados.

En cuanto al tipo de desviación más frecuente tenemos en primer lugar la desviación que refiere a hacer regular las formas irregulares, le sigue la que implica una irregularidad distinta a la estándar, y luego la que agrupa a las formas que exhiben una asignación incorrecta de desinencia. En último lugar hallamos la desviación que alude al hecho de hacer irregular una forma que es regular.

### Evaluación de hipótesis en el aspecto informático

La hipótesis es que se puede lograr el análisis automático de las formas verbales irregulares desviadas si se adapta la formalización lingüística declarada en los programas informáticos Smorph y Nooj que fueron diseñados para el reconocimiento de lenguas naturales.

Además, que el sistema Nooj ofrece mayores posibilidades que Smorph en lo que refiere al reconocimiento de formas desviadas propias de la interlengua de aprendientes de español.



Luego de realizar modificaciones en el archivo Entradas, Rasgos y Modelos en Smorph fue posible el reconocimiento automático de desviaciones 1 y 2 pero alterando cada uno de los modelos de los verbos irregulares para agregar las formas desviadas y además creando verbos inexistentes en español en el archivo Entradas.

Para las desviaciones del tipo 3, debieron ingresarse en el archivo Entradas una a una todas las ocurrencias, generando esto no sólo una sobrecarga de este archivo sino implicando un recorrido inverso: ingresar primero los ítems que voy a encontrar en el texto. Es decir, no satisface la exigencia de cobertura de los casos desviados.

Con respecto a las desviaciones tipo 4 no fue factible reconocerlas ya que podrían detectarse mediante las terminaciones verbales pero en este caso fallaría la precisión ya que no distinguiría entre verbos del español y verbos de la interlengua, los etiquetaría de la misma forma.

Nooj logró detectar y analizar los casos de desviaciones 1, 2 y 4 mediante la generación de gramáticas productivas que funcionan aplicando diccionarios del módulo español, distinguiendo y etiquetando a las formas desviadas como verbos de interlengua.

Además pudo proporcionar una etiqueta (RI) o (RR) con el fin de indicar que debe emplearse la raíz irregular o regular de ese verbo para formar un verbo irregular que pertenezca al español estándar.

En resumen, el sistema Nooj no sólo permitió realizar el reconocimiento automático de una manera mucho más económica, sino que además proporcionó una mayor cobertura, abarcando los casos desv4, y el 65% de desv3 a partir de la creación de una gramática morfológica que reconoce que cualquier conjunto de caracteres más desinencias verbales es un verbo de interlengua.

## **Aportes de mi tesis**

La investigación ha realizado una contribución al estudio de la lingüística computacional en lo que refiere al análisis automático de textos de interlengua mediante el empleo de herramientas informáticas que fueron creadas para el análisis de lenguas naturales. El objetivo pudo cumplirse en la medida que fueron efectuadas adaptaciones y se generaron gramáticas productivas.

Página | 170

Cabe aclarar que el estudio se sirvió de la labor realizada por el grupo INFOSUR respecto de la implantación en máquina de todas las categorías de palabras y de modelos flexivos en el software Smorph para el tratamiento automático de la lengua española.

Por lo tanto, considero que abre el camino a futuras investigaciones sobre análisis automático de la interlengua mostrando las posibilidades que ofrece el uso de herramientas informáticas como recurso didáctico para la enseñanza-aprendizaje del español como segunda lengua.

En cuanto al ámbito de la adquisición contribuye al conocimiento de la morfología verbal irregular y del origen de las desviaciones, evaluando de manera positiva la hipótesis de que estas formas idiosincrásicas responden a una causa intralingüística.

## **Proyecciones**

La presente investigación se propone abrir el campo de estudio en lo que respecta al tratamiento informático de la interlengua, mostrando las posibilidades que ofrece el sistema Nooj para ser empleado como recurso didáctico en el proceso de enseñanza-aprendizaje del español como L2, teniendo en cuenta que el propio estudiante de español puede analizar sus producciones y obtener una corrección sobre las formas verbales desviadas.

Aït- Mokhtar Salah y Rodrigo Mateos José Lázaro (1995) "Segmentación y análisis morfológico de textos en español utilizando el sistema SMORPH. SEPLN, 17, 29-41

Alcoba, Santiago (1999) "La flexión verbal" en Gramática Descriptiva de la lengua española, dirigida por Bosque, I. y Demonte, V. Espasa, Madrid.

Alexopoulou, A.(2005) "Errores intralinguales e interlinguales en la interlengua escrita de aprendientes griegos deE/LE". Actas XVI Congreso Ai Centro Virtual Cervantes.

Alexopoulou, A. (2010) "La función de la interlengua en el aprendizaje de lenguas extranjeras". Revista Nebrija de Lingüística aplicada 2010, 9.

Archibald, J. (Ed.). (2000) *Second Language Acquisition and Linguistic Theory*. Massachussets: Brasil Blackwell.

Baralo Ottonello M. (2004). "La interlengua del hablante no nativo" En Sánchez Lobato, J. Santos Gargallo, I. y Pinilla, R. (Eds.) *Vademécum: enseñanza y aprendizaje del español como segunda lengua y como lengua extranjera*. (pp. 369 – 389). Madrid: SGEL.

Baralo Ottonello M. (2004a). *La adquisición del español como lengua extranjera*. Arco Libros, Madrid.

Beltrán, C. (2009) *Modelización lingüística e información estadística en el análisis automático de textos*. (Tesis de Doctorado). Ediciones Juglaría, Rosario.

Beltrán, C. (2010) Estudio y comparación de distintos tipos de textos académicos. Biometría y Filosofía.Revista de Epistemología y Ciencias humanas Nº 2, 130-139, Grupo IANUS, UNR, UNL, Rosario.

Blanco, Picado A. (2012) "El error en el proceso de aprendizaje en Cuadernos Cervantes". Disponible en [http://www.cuadernos cervantes.com/art\\_38\\_error.html](http://www.cuadernos cervantes.com/art_38_error.html)

Bonino, R. (2008) *Análisis automático de las construcciones causativas*. Tesis doctoral. Facultad de Humanidades y Artes, UNR

Bonino, R. (2009) "Presentación de la lingüística computacional" en La interlengua de los aprendientes del español como L2. Aportes de la lingüística informática. Centro de adquisición de estudios del lenguaje. Juglaría, Rosario.

- Bonino, R. (2011) "Una propuesta para la implantación de la morfología verbal del español" en Revista *INFOSUR* Nº 5, págs. 79-86.
- Brucart J. (2005) "La gramática en ELE y la teoría lingüística: coincidencias y discrepancias", RedELE (Revista electrónica de didáctica del español como lengua extranjera), número 3, disponible en [RedEle/Revista/2005\\_03/2005\\_redELE\\_3\\_07Brucart.pdf?documentId=0901e72b80e063f1](http://RedEle/Revista/2005_03/2005_redELE_3_07Brucart.pdf?documentId=0901e72b80e063f1)
- Chomsky N.(1989) *El conocimiento del lenguaje: su naturaleza, origen y uso*. Alianza Editorial.
- C.V.C.(Centro Virtual Cervantes) Diccionario de términos clave de ELE. [http://cvc.cervantes.es/ensenanza/biblioteca\\_ele/diccio\\_ele/diccionario/transfereencia](http://cvc.cervantes.es/ensenanza/biblioteca_ele/diccio_ele/diccionario/transfereencia).
- Corder (1967) "The significance of Learner's Errors" en *International Review of Applied Linguistics*, num 5.4. pp.161-170
- De Alba Quiñones V. (2009) El análisis de errores en el campo del español. Revista Nebrija. Disponible:[https://www.nebrija.com/revistalinguistica/files/articulosPDF/articulo\\_5316e9a63dc81.pdf](https://www.nebrija.com/revistalinguistica/files/articulosPDF/articulo_5316e9a63dc81.pdf)
- Dister A. (1997) Problematique de fins de prhase en traitement automatique du français, en Á qui appartient la ponctuation? Actes du colloque international et interdisciplinaire de Liège. Bruxelles.
- Ferris, D. (2005). *Treatment of Error in Second Language Student Writing*. U.S.: The University of Michigan Press.
- Guillot, D. "Lingüística computacional: del prototipo a la aplicación". Publicación de Facultad de Filosofía y Letras de la Universidad Nacional de Cuyo. JALIMI 2005.
- Hipperdinger Y. (2006). *Estudios sobre contacto inter- e intralingüístico*.Universidad Nacional del Sur. Departamento de Humanidades, Bahía Blanca.
- Lado, R. (1957) *Linguistics Across Cultures*, Ann Arbor, University of Michigan Press
- Larsen-Freeman, D. & Long, M. (1984): *Introducción al estudio de la adquisición de segundas lenguas*.Madrid: Ed. Gredos.
- Lavid, J. (2005) Lenguaje y nuevas tecnologías. Nuevas perspectivas, métodos y herramientas para el lingüista del siglo XXI. Cátedra. Madrid.

Liceras, J. (1992) *La adquisición de las lenguas extranjeras: hacia un modelo de análisis de la interlengua*, Madrid, Visor

Liceras, J. And L. Díaz (2000). Triggers in L2 acquisition: the case of Spanish N-N compounds. *Studia Linguistica* 54: 197-211.

Marco Común de Referencia europeo. Disponible en [http://cvc.cervantes.es/ensenanza/biblioteca\\_ele/marco/cvc\\_mer.pdf](http://cvc.cervantes.es/ensenanza/biblioteca_ele/marco/cvc_mer.pdf)

Martín J. (2004) “La adquisición de la lengua materna (L1) y el aprendizaje de una segunda lengua (lengua extranjera): procesos cognitivos y factores condicionantes”, en SÁNCHEZ LOBATO J. /SANTOS GALLARDO I. (dir) *Vademécum para la formación de profesores*. Madrid, SGEL, PP. 261-286.

Martín Vegas, R. (2014) *Los verbos irregulares del español*. Biblioteca Nueva, Madrid

Martínez A. La hipótesis de la transferencia de la L2 y la adquisición del orden de palabras.

Mendenhall, W. Wackerly, D. Scheaffer, R. (2002) *Estadística matemática con aplicaciones*.

Múgica, Nora (2016) *Morfología de los tiempos verbales del español. LA FLEXIÓN VERBAL. REVISTA DE GRAMÁTICA DEL ESPAÑOL PARA EXTRANJEROS. DISPONIBLE EN WWW.REGRAELE.COM.AR*

Nemser W. (1971) “ Approximative systems of foreign language learners” en IRAL International Review of Applied Linguistics in Language Teaching 3 , págs. ++/+, - SAN

Real Academia Española (Asociación de Academias de la lengua española) (2009) *Nueva Gramática de la lengua española*. Madrid, Espasa.

Reynar, J. y Ratnaparkhy, A. (1997) A Maximum Entropy Approach to Identifying Sentence Boundaries, en ANLP 1997; arXiv:cmp-lg/9704002 v1.

Rigamonti Daniela (2006) *Problemas de lingüística de la adquisición y enseñanza del E/Le a italófonos*. MPW MOTS PALABRAS WORDS. Disponible en <http://www.ledonline.it/ledonline/rigamonti/Rigamontiproblemas.pdf>

Rivero Silvia (2006) *Adquisición de segunda lengua*. Centro de Estudios Adquisición del Lenguaje (UNR), Juglaría, Venado Tuerto, Argentina.

Rodrigo A. (2010) Tratamiento automático de textos: el sintagma adverbial núcleo. Tesis doctoral. Universidad Nacional de Rosario.

Santos Gargallo Isabel (2010) *Análisis contrastivo, análisis de errores e interlengua en el marco de la lingüística contrastiva*. Editorial Síntesis.

Selinker, Larry (1972) "Interlanguage", en *International Review of Applied Linguistics*, num 10.3 pp.209-31.

Silberztein, M. (2003): *Nooj Manual*, [en línea], <<http://www.nooj4nlp.net/NooJManual.pdf>>

Silberztein, M. (2015) *La formalization des langues: l'approche de Nooj*: ISTE eds. London.

Solana Zulema y otros (2006) *Recursos informáticos para el tratamiento lingüístico de textos*. Centro de Estudios de Adquisición del Lenguaje. Facultad de humanidades y Artes. UNR. Ediciones Juglaría.

Solana Zulema y otros (2006) *Morfología del verbo español*. Colección Infosur. Centro de Estudios de Adquisición del Lenguaje. Facultad de humanidades y Artes. UNR. Ediciones Juglaría.

Solana, Zulema; Beltrán, Celina; Tramallino Carolina (2009) "La implantación en máquina de la interlengua de los aprendientes de español como L2: los sufijos formadores de nombres" (en colaboración) (2009) en *La interlengua de los aprendientes de español como L2. Aportes de la lingüística informática*. Editor: Centro de Estudios de Adquisición del Lenguaje. U.N.R. Ediciones Juglaría, Rosario.

Solana, Zulema; Beltrán, Celina; Tramallino Carolina (2013) "Análisis automático morfológico con las herramientas SMORPH Y NOOJ" (En *Revista de epistemología y ciencias humanas*. Grupo IANUS, Docentes de la Universidad Nacional de Rosario y la Universidad del Litoral. ISSN 1852-625X. Nº 5, abril de 2013.

Solana Z., Tramallino C., Méndez B. y otros (2014) *Estudios del lenguaje: niveles de representación lingüística* (Proyecto Latin. Iniciativa latinoamericana de escritura de libros abiertos. Disponible en [http://www.latinproject.org/books/Estudios\\_Lenguaje\\_CC\\_BY-SA\\_3.0.pdf](http://www.latinproject.org/books/Estudios_Lenguaje_CC_BY-SA_3.0.pdf)

Sonsoles Fernández (1995) "Errores e interlengua en el aprendizaje del español como segunda lengua" en *Didáctica*, 7, 203-216, Servicio de Publicaciones UCM, Madrid

Torruella J. Llisterri J. (1999) "Diseño de corpus textuales y orales" en BLECUA J. M., CLAVERÍA G. SÁNCHEZ C. TORUELLA, J. (editores) . *Filología e informática*, Univ. Autónoma de Barcelona, Editorial Milenio, pp 45-77.

Tramallino, C. (2009) "Formas verbales irregulares en la interlengua de aprendientes de español como L2" en *La interlengua de los aprendientes de español como L2. Aportes de la lingüística informática*. Editor: Centro de Estudios de Adquisición del Lenguaje. U.N.R. Ediciones Juglaría, Rosario.

Tramallino, C. (2010) "Construcción de formas verbales en la interlengua de aprendientes de español como L2. Tratamiento automático para participios y gerundios irregulares." en *Revista de epistemología y ciencias humanas*. Grupo IANUS, Docentes de la Universidad Nacional de Rosario y la Universidad Nacional del Litoral. ISSN 1852-625X. Nº 3, abril de 2010.

Tramallino, C. (2011) "Verbos irregulares utilizados por aprendientes de español. Técnica estadística no paramétrica en la obtención de corpus de diferentes lenguas" en *Revista INFOSUR* Nº 5, págs. 41-52. ISSN 1851-1996, octubre 2011.

Tramallino, C. (2012) "Análisis automático de la interlengua de aprendientes de español: la presencia de determinantes indefinidos en el sintagma nominal núcleo" en *Revista INFOSUR* Nº 6, págs. 75-87. ISSN 1851-1996, diciembre 2012.

Tramallino, C. (2013) "Análisis morfológico con herramientas informáticas. Reconocimiento de nombres en textos de español con el sistema Nooj" en *Revista Lingüística y Literatura* ISSN 0120-5587. Universidad de Antioquía, Medellín, Colombia. Año 34, Nro 64 (julio-diciembre 2013) ISSN 01205587. Disponible en <http://aprendeonline.udea.edu.co/revistas/index.php/lyl>

Wayne W Daniel, (1999): *Bioestadística*. Limusa, México

Vázquez A. (2007) "Análisis de errores, el concepto de corrección y el desarrollo de la autonomía" II Jornadas y I Congreso Internacional de Enseñanza de Español como Lengua Extranjera organizadas por el Centro de Estudios del Español como lengua extranjera de la UNR.

## ANEXOS

### Anexo Nº 1: CORPUS

#### 1.1. Corpus inglés

Página | 176

(Selección de textos)

Podría decir que mis gustos son similar de los gustos de Gaby pero no me gusta el tenis. Soy un fanática del fútbol tambien. Me gusta tomer helados y me gusta passar tiempo con mi amigas. Pero no me gusta la tecnologia y las paginas web.

Gaby se gusta hacer páginas web. Me gusta helado también pero no como helados porque voy a “engordar”.

A me mi gustan tocar el piano. Tengo uno a mi casa. Me también me gustan los deportes como el jockey sobre hielo y football americano. Estan de depuortes muy populares en Quebel.

A mi no me gusta los deportes pero envio las personas que le gusta por qué los deportes son bueno para la salud. Yo no quiero mucho los animales y me gusta mas el viernes.

Prefiero me preocupar de los derechos humanes, antes los derechos de los animales.

No quiero creadar un club de protección de animales en extinción.

¿Cuál días es disponible? ¿Cuán costar con apart? ¿Cómo funcionar para reserva?

Yo apprenda tu quieres venir a Rosario con una amiga. El Appart hotel es un hotel muy bella y el persone es muy simpático. Si tu quieres haber information, puedes telefonar.



Yo envio me fotos a la CLEC in quiero obtener vos fotos tambien para hacer una "album foto" de notra viajar a Rosario. Yo mucho quiero el viajar y muchas gracias para el ayudar del CLEC.

Pedir de información

Envío las fotos del fiesta de bienvenida de la universidad de Rosrario. Quiero hacer vos fotos también.

Information por dos amigas

Te escribo por que nessecito information sobre la disponibilidad del apart y las tarifas tambien. Tengo dos amigas que van en la Argentina por vacaciones.

Soy de Canadá y quiero saber si es posible de ir a vos hotel. Soy con una amiga y queremos venir el último fin de semana de septiembre. Somos muy interesante a visitar Rosario y tambien a vivir en su hotel.

Te escribo para te enviar fotografias tomadas durante nuestro estadia en Rosario. Quiero saber si puedes enviar las fotos que tú y tus amigos toman cuando vamos en un bar.

Soy de regreso en Montreal. ¡Soy muy dicha de ver mi familia! Me gusta racontar mi viaje y quiero dar fotos. Pensio qué durante la fiesta de bienvenida tu pones de nosotros. ¿Es possible que tu enviades las fotos? Me encanta esta fiesta y quiero un recuerdo.

He venido a Rosario el mes pasado. Cuándo hemos hecho la fiesta el 10 de junio, vos haces una fotografía de nosotros. Es posible de me le enviar? Quiero también que envíe las fotos para todas las personas del grupo. Muchas gracias para todas que hacía para nosotros. Me gusta mucho Rosario y me gusta mucho todo que hacemos en nos viaje.

Queremos una habitación doble y los desayunos incluidos. Puede responderle más rápidamente por favor porque si es demasiado caro, tenemos buscar otra opción.

Somos dos canadienses y queremos pasar una fin de semana en Rosario. ¿Es posible de hacer una reservación? ¿Cuánto cuesta para un departamento? ¿Puede escribirme como podemos reservar? Esperamos que vamos a encantar Rosario como encantamos la Argentina.

Mucho gracias.

Querido Katie,

¡Yo he hecho tanto! He conocido mi familia de acogida y los me gustan mucho. Son muy agradables y yo creo que me le gusta también. ¡Fui a la escuela ayer y los uniformes son muy muy muy feo!

Vamos de compras mañana y espero que yo pueda comprar algo de ropa de Argentina. La ropa es bonita y quiero que algunos. Argentina es increíble y lo me gusta ya. ¡No puedo esperar para hacerlo todo! Te extraño mucho y no puedo esperar a verte.

Adiós.

Hola Laura! Como estáis?

Aquí es fantastico! He conoscado muchas personas sympaticas y mi nueva familia es super!

Voy al instituto con amigas y los profesores son sympaticos, pero no comprendo mucho...

Qué lástima, por que queria entender que hablan! Pero mis amigas me teache bien y paulatino aprendo la lengua.

La primera semana en Rosario ha estido super, hamos visitado mucho de la ciudad y hamos aprendido español en un curso con jueves del mundo! Ahora son buenos amigos de mi;-)

He vestido mucho del país y quiero ir a Buenos Aires por que es la capital de Argentina y es grande.

Aquí en Arrecifes no han un club de atletismo pero quiero bailar Tango en una escuela con amigas.

Son una pequeña grupa pero sería muy bien!!!

Mi nueva hermana Delfi es super sympatico y compramos bien! Dormimos en una calle así en las noches no dormimos mucho;-)

Espero que tu vamos bien en Inglaterra y que ti familia es bien tambien!!!

Espero tu respuesta!

Hola mis amigas! Quiero ver ustedes mas pero necesito volver a los Estados Unidos. Sería bueno que me tomaste a las discotecas y la ciudad de Mendoza porque me lo guste. Es una pena que tengo que salir y lo muy siento. Me encantaria hablar contigo a casa por la telefono o la computadora. Es una buena idea? Muchos gracias para el tiempo que me di.

Hola mis amigas! Quiero ver ustedes mas pero necesito volver a los Estados Unidos. Sería bueno que me tomaste a las discotecas y la ciudad de Mendoza porque me lo guste. Es una pena que tengo que salir y lo muy siento. Me encantaria hablar contigo a casa por la telefono o la computadora. Es una buena idea? Muchos gracias para el tiempo que me di.

Estoy bien porque viaje a las montañas con mi familia este semana. Fuemos al pueblo de Inlet Forge donde vi muchos animales. Mi escuela termino la semana pasado. Muchos del dios voy al cine con mi novio o camino en el parque. Me gusta caminar mucho porque necesito el ejercicio. Tambien hago yoga con solamente mi porque la silencia esta muy tranquila. La moda es mi favorita campo de estudio. Tambien quiero estudiar en Australia porque me encanto viajar. Espero que hablamos pronto! Adios.

Saludo. Me llamo Martin Galván y estuve un pasajero en un vuelo proveniente de Madrid. Estoy indignado porque sufrí inconvenientes. Tuve una demora de dos horas en Barajas. También, debería haber mas bebida y comida y el vuelo seria mejor con auriculares. Quiero un reembolso. Espero una pronta solucion.

Martin Galván

Estoy del acuerdo con Doctor Jorge porque yo creo que nosotros necesitamos menos estrés en nuestras vidas.

Creo antes, gente comia y preparaba comida mas sana que ahora.

Antes, la comida preparaba en la cocina de la family, ahora, la comida es delivery.

Ahora, se comen pizza, empanadas, panchos, hamburguesas y papas fritas.

Antes, comian platos casero.

Antes, la comida era casero, hoy la es delivery y Fast food.

Hoy, gente tiene mas problemas de salud como obesidad, diabetes, y trastornos cardiovasculares.

El cambio afecta los niños fundamentalmente.

Familias no tienen tiempo a comer a la mesa.

Es necesario recuperar la tradición de la cena en familia y que la comida sea casera.

Hoy, yo como mucha pan y frutas.

En los Estados Unidos, comia muchas verduras con tofu.

Mimadre o yo preparamos la comida.

Cocinabamos arroz integral con tofu y verduras

1. En mi relato yo creía que Pablo quiere Marisa mucho y en este Pablo la quiere mucho
2. En mi relato hacía mucho que no se veían porque se pelearon cuando Marisa se casó y en este no se pelearon.
3. Yo creía que se conocieron de escuela, estudiaban juntos. Es verdad.
4. Yo pensé que tenían una relación porque eran buenos amigos pero eran en enamorado.
5. Yo supuse que Pablo no se casó pero se casó pero su matrimonio no funcionó y se divorció.

Marisa quería divorciarse de Gustavo, y necesitaba saber si él aún la amaba. Gustavo dijo que la amaba y se fueron a vivir juntos.

Pablo y María eran buenos amigos y eran en enamorado. Estudiaban en la misma escuela. Hace tantos años que no se ven porque desde ella se casó su vida cambió mucho. Pablo se casó pero su matrimonio no funcionó y se divorció.

José de San Martín nació el 25 de febrero de 1778 en Yapeyú. Su madre se llamaba Gregoria Matorras, y su padre, Juan de San Martín. Tenía 4 hermanos. Se casaba con María de los Remedios de Escalada. Ellos tenían una hija, se llamaba Mercedes. En 1784, él llegó a España con su familia. Desde 1789 hasta 1811, ingresó al Regimiento de Murcia. En marzo de 1812, llegó a Buenos Aires y creó el Regimiento de Granaderos. En 1813, hizo su primera intervención. En 1817, cruzó la Cordillera de los Andes. En 1818, declaró la Independencia de Chile. En 1821, declaró la independencia en Perú.

Querido Gus,

Ahora en Rosario, hay un festival de jazz grande. Quieres ir a la Sala Cultural Mano a Mano conmigo el viernes 5 a las 8:30 en la tarde? Hay va a ser una presentación de Los Cornalitos de Jazz y Rock. Muy bueno!

Chau,

Compre los ingredientes. Coloque en una caserola, 5 cucharadas de grasa hasta que se derrita, agregue la cebolla picada, dore unos momentos y ponga la carne (antes piqela a mano pasela por agua hirviendo. Condimente con sal. Agregue un chorro de agua fria y retire de fuego. Mezcle las papas picadas y cocidas. Prepare este relleno la noche anterior. Ponga en un recipiente y deje enfriar. Deje enfriar el relleno para que armar las empanadas resulte fácil. Sobre cada uno de los discos de masa coloque 1 cucharada de relleno, junte los extremos y “repulgar”. Antes de cerrar las empanadas, coloque sobre el relleno cebolla verde y los huevos duros (picelos chicos). Free en una cacerola con grasa caliente bien la grasa.

Rose y Gabriela ya comieron asado y tomaron mate. Todavía no bailaron tango. Todavía fueron a la cancha pero ya probaron el dulce de leche. Ellos todavía no fueron a la Patagonia pero ya visitaron el Noroeste. Todavía no conocieron las cataratas y todavía no sacieron una foto en el obelisco.

Ya comí asado y ya tomé mate. Todavía no bailé tango ni fui a la cancha. Ya probé el dulce de leche y ya fui a la Patagonia. Todavía no visité el Noroeste ni concí las cataratas. Ya sacé una foto en el Monumento a la Bandera



Hola Valerie!

Soy en Argentina. Vivo en Rosario, una ciudad que se situa 300 kilometros al lado de la capital Buenos Aires. Me gusta mucho los rosarinos. Son muy simpáticos personas. La comida argentina es bueno. Como ravioles con salsa mixta y ensalada. Visito la ciudad un pocito. Hablo español con cuatro personas. Comenco los cursos de español en la mañana. Los habitaciones son muy comfortable. Espero tu respuesta.

Te quiero

¡Hola chica! Te escribo de Argentina, estoy en la ciudad de Rosario. El viaje fue un poco difícil, pero estoy muy contenta de estar aquí con los otros estudiantes del colegio. No te vas a creer esto: ¡hace frio en Argentina! Es que ahora, es el hivierno en america del Sur.

Ayer, yo he visto toda la ciudad de Rosario en un autobús para los teuistos. Yo puedo decir que está interesante pero mi corazón está en Montreal, te lo juro. Esperando que tu salud está bien, muchos besos.

¡hola!

Fui en el Argentina de el 25 de mayo hasta el 19 de junio. Quiero agradecer para la recepción el la ciudad Rosario que me gusta, mucha. ¿Quiero hacer las fotos que vos tienes de la fiesta de bienvenia del viernes 30 de mayo? Me necesita porque no tengo foto esta día. Me gusta mucho todas días en Rosario 4 espero que vos gustan nosotros tambien.

Querido Gauthier:

Espero que las vacaciones te hayan ido bien. Te escribo ese mail para informarte de un evento cultural muy original: la segunda gran liberación de libros. El concepto es muy simple: deja tu libro preferido en un lugar público, la persona que lo encuentre lo lee y lo deja de vuelta para una otra persona. Asi se forma una cadena. ¡El evento empieza mañana! Tendrían que participar porque siempre me dijiste que querias compartir tus libros con desconocidos. Además, leer es una actividad bastante solitaria, pero esta oportunidad te permitiría de en contra gente con mismos intereses. Y vos también podés encontrar un libro de un desconocido, es una experiencia bastante excitante.

Si quieres más información, llamame o anda sobre [www.librolibreargentino.com](http://www.librolibreargentino.com)

Me ordenó que pidiera al mozo que me trayera la comida. Me dijo que Juan llegaría martes. Le dije que esperaba que lleguese solo.

Le preguntó a la señora si quería que le trayese el te y la señora respondió que le rueguía que se lo pondiera azúcar.

Dijo que en esa empresa les pagaban muy mal y que los obreros exigían que les pagueran mejor.

Me dijo que cada tanto escribía un poema y yo le dije que le suplicaba que escribiera uno para el cumple de la abuela.

Me preguntaron si podían alquilar un video y yo les dije que les prohibía que se alquilaran uno.

Le dije que necesitaba que el redactor tuviera experiencia. Me sugirió que me apuriera

Se recomienda no entrar al agua apenas de terminar de comer. Se sugiere no dejar al niño inacompañado al agua ni a ningún de la playa ni lejos de la playa.

No te olvides de llevar mucha plata para comprar ropa y otros objetos coyas que son típicos del Noroeste argentino.

Ustedes todos saben de mi situación. Estoy en 5to y no tengo idea que quiero estudiar. Ayer hablé con mis padres, ellos me aconsejaron que estudiara política porque siempre me sacé buenas notas pero yo les dije que no haya querido ser político. Como otra oportunidad ellos surgieron que haya seguido con la idea de estudiar medicina.

A mí también me gusta mucho la filosofía, entonces proponí que estudiara eso.

Pero mis padres me dijeron que fuera difícil conseguir un trabajo filosófico. Ellos me recomendaron eligiera psicología sea posible pero ya no lo sé si lo haría.

### Viaje al Noroeste

Cuando llegé al terminal del ómnibus en Rosario casi nadie era allá entonces tenía que esperar. Con un poco de tiempo algunos conocidos llegaron y nosotros empezamos a hablar. Acá todos los estudiantes tienen algo para contar.

A las 9 la viaja empezó que suerte que algunos chicos trajeron mate, así la viaja no era tan aburrido.

Viajábamos toda la noche y en aquella noche dormí pocas horas, todo el tiempo algunos chicos hablaron y también porque en el colectivo hice mucho frío.

Entonces estube muy alegre cuando el colectivo llegó al terminal de Salta.

Todas las excursiones estubieron tan lindas y la naturaleza... Yo nunca había visto algo tan hermoso. A la primera noche fuimos a una peña y en esa peña comimos comida típica del norte de Argentina, en frente bailadores de folklore bailaron ¡una show de baile muy copada!

El penúltimo día fuimos a una excursión que llegaba todo el día

La última parada de esta excursión fue la salina en los Andes un lugar tan surreal y increíble. Nunca jamás pude imaginar tanta sal.

Cuando el colectivo a Rosario llegó al terminal me sentí muy melancólico. Me senté en el colectivo pero en aquel momento quería quedarme en Salta uno de los lugares más lindos que hay en el mundo...

Mis pensamientos fueron interrumpidos por la voz de una chica. Me preguntó si se podía sentarse a mi lado. Tuve vergüenza de preguntarle cómo se llamaba. Discutaba por todo, pedí dos porciones de pizza y ella quería ir al baño.

A esa hora tuve que irse. Estoy muy contento que la conocí

#### Presistencia de Galtieri

Al finalizar los 5 años de la presencia de Videla éste fue reemplacado por en general Roberto Eduardo Viola. No logró conseguir una apoyo consistente. La corta duración de la preesistencia de Viola dio a conocer los conflictos internos del regime militar.

Los señores pideron y los mozos traeron la comida. Luego venieron los coordinadors y después haciste el bolso para volver.

A Carolina *se gusta* las novelas ¿y a vos?

A me, me gusta las novelas tambien.

¿Vos prefereís las películas románticas o de acción?

Yo preferio las películas de acción.

A mi mejor amiga se gusta la música pop.

Cuando termina de trabajar, a las siete de la tarde vuelve a su casa, se cambia de ropa y se al gimnasio. Los lunes y miércoles, después del gimnasio, juega al tenis con su amigo. Los martes y jueves salí a cenar con su novia.

Patricia estaba sentando en su sillón en el dormitorio casi dormiendo ella escuchó una persona que gritó afuera miro du su ventana pero no pude ver nada. Ahora no tenía más sueño se sintió nuevamente en la sillón y empezó a leer... y un grito tan fuerte como Peter nunca lo escuchó fue afuera... el miró de la ventana pero no vío nada. Empezó a tener panique no sabía que ella pude hacer. Patricia se dio vuelta pero no había nadien.

Hasta ahora nadie ha escrito una instrucción para pelear una banana. En primiero tenés que sacar una banana. Despues tiene que sacar este parte de cascara, la mano derecha pusque otra parte de la cascara en el pico de la banana. Este acción se repite hasta que el punto de la banana hasta la mitad, en que se equentre la otra mano, esta libre de cascara. Ahora se puedo manejar la banana al la boca pero antes de comerla es muy nesecario de abrir la boca.

### CORPUS PORTUGUÉS (Textos 7 al 29)

En el cuadro tiem un grupo de personas que viven en una región distante, que se encuentran ubicadas para hacer muchas cosas como: cantar, charlar, calentarse del frío alrededor del fuego en la cabaña (enquanto)mientras una senhora lava platos.

El pintor Molina Campos describe una casa con muchas personas. La casa és pobre e las personas están en vuelta de una hoguera porqué está mucho frío. Ha varios hombres en la cabaña . Ellos son personas con mucho años y mui graciosos. Ellos están hablando. Ha también un brase porque el tiempo está mui frio. Esas hombres

. El artista retrata la imagen de una casa de campo en na cual una familia humilde si agrupan en redor de una hoguera para acalentarse del frio. Algunas figuradas retratadas tales como: cajones, escoba, y panelas fijadas en las pareds, **son** muy interessants, pues demonstnan la simplicidad de la vida campestre incluyndo en las

. El artista retrata la imagen de una casa de campo en na cual una familia humilde si agrupan en redor de una hoguera para acalentarse del frio. Algunas figuradas retratadas tales como: cajones, escoba, y panelas fijadas en las pareds, **son** muy interessants, pues demonstnan la simplicidad de la vida campestre incluyndo en las personas. La paisaje **es** agradable y muy común en la realidad de países como Brasil e Argentina.

. El artista retrata la imagen de una casa de campo en la cual una familia humilde se agrupan en redor de una hoguera para calentarse del frío. Algunas figuras retratadas tales como: cajones, escoba, y panelas fijadas en las paredes, **son** muy interesantes, pues demuestran la simplicidad de la vida campestre incluyendo en las personas. El paisaje **es** agradable y muy común en la realidad de países como Brasil e Argentina.

La casa del interior, donde se encuentran muchos amigos hablando. Parece una familia, incluyendo a un animal. Están sentados, alrededor de una fogata.

El autor retrata en su imagen un local donde se encuentran personas que **están** sentadas en torno al fuego. La casa **es** muy rústica y el ambiente **es** de mucha tranquilidad.

Es una noche y en una cabaña existen algunas personas al redor de una fogata hablando mientras toman mate. **Es** un grupo formado por hombres adultos mayores. **Hay** también un perro y un gato, en el fondo de la única mujer en el local. **Es** una casa sencilla, utilizan cajones para sentarse.

El cuadro muestra la casa con personas en el interior festejando algo a mayoría sentada y todos felices. Hay una persona, debe ser una mujer preparó bebida, el o que parecer. Están ubicados al fondo el cantor historias. El fuego calienta el ambiente, debería estar muy frío.

Cuando nosotros llegamos a Rosario estaba muy frío. Temo que nos congeláramos. Rosario es una ciudad muy hermosa con calles arbolizadas, en el frío es más bella. Las personas mayores andan solas por las calles y cuando las miro recuerdo a mi mamá y papá. Estoy con mucha nostalgia de ellos. Las casas acá están muy caras, más que en Buenos Aires. Estoy en la misma hostería con localización distante de la universidad. Ya estoy acostumbrada a la medialuna, pues en la mañana es en toda mañana.

Día 19 de julio 11:30 hs desembarquei en la aeropuerto de Ezeza, 13:30 embarquei en un ómnibus para la ciudad de Rosario, lleguei 18:00 hs en la hotel Plaza, la temperatura mucho frio, estaba aproximadamente nuevo grados. Venir para concluir la segunda módulo del curso de Doctorado en Ciencias la Educación por la UNR. Encuentrei la ciudad linda como siempre, más mucho frío. Para mi sorpresa lo edificio de la UNR donde teníamos concluida la primera módulo encontrose en reforma, por la tanto ficamos más distante del hotel donde estamos. Lo nuevo edificio de la UNR localizase la rua Corrientes. Entonces Encuentrei entusiasmado por el curso, la dificultad que encuentro es la pronuncio, hasta lo término de este curso de español estarei fuerte en la defeza de mi tesis.

Cuando estibe en Rosario por la premera vez, sintió una maravillosa sensación de que algo bueno iba a suceder como: conocer gente nueva, hacer muchos amigos, construir nuevos conocimientos, y fue realmente lo que pasó. Sin embargo, el impacto al regresar a Rosario después de 6 meses me sentí la misma alegría. El impacto de la búsqueda de una bela ciudad rodeada por una fria y oscura era sorprendente. Por otro lado, volver a conectar con viejos maestros, amigos de otras provincias, así como nuevos profesores fue una alegría para todios.

Ahora en julio llegué a Rosario a la noche. Pronto percebi que había diferencias. La primera fue la necesidad de utilizar ropa grossa para protegerme del frío. Hubo otro cambio, las clases se encuentran en otro lugar. Tenemos que hacer una larga caminada, un poco más de veinte minutos. Reencontramos a los maestros, los amigos y sentimos la ausencia de otros. Como Rosario mostró que algo ha si cambiado en su paisaje, paréceme que nosotros también, pero en nuestra singularidad.



Cuando llege en Rosario, 16 de enero de 2012, me sentí sorprendido con las novedades. El tiempo estaba muy agradable. Muchas cosas eran diferentes de mi ciudad. Había muchos perros por las calles de la ciudad. También había muchas parejas de viejos por toda la ciudad. Sentí muchas diferencias y las personas hablaban muy poco. Ellos casi no sonríen, parecían no gustar mucho de los brasileños. Ahora, julio de 2012, por segunda ocasión, me sentí más sereno. Vi muy pocos perros por las calles. La ciudad parecía mucho más bella pero el tiempo muy frío. Los profesores parecen más satisfechos. También el edificio de la universidad está más confortable. La ciudad parece que está con mucho más personas. Cada vez estoy sintiéndome como un rosalino. Estoy sintiéndome en mi propia casa. Cuando cese mi doctorado, quiero volver otras veces en Rosario.

Yo moro en Brasil y estudio en Rosario. Yo empecé mi curso de Doctorado en enero, cuando llegué en la ciudad de Rosario el tiempo estaba caliente como en Brasil. Ahora en julio el tiempo está muy frío. Nosotros estamos estudiando filosofía, instituciones e historia, también estamos estudiando el español. La mi profesora se llama Carolina, ella es una buena profesora y enseña muy bien. Estamos estudiando los pronombres demostrativos, los posesivos e también la voz pasiva. Me gusta mucho estudiar otra lengua.

Rosario hoy está más fría que en enero cuando estuve acá. Las plantas están con las hojas secas y sin belleza. A la noche las personas se encuentran para beber vino y conversar sobre la familia, el trabajo y otras cosas más. La impresión que se tiene es de una ciudad que para en el tiempo para ver el tiempo pasar. Todo está más frío y las noches de tango en la plaza no acontecen y eso deja la ciudad menos caliente. Entretanto, Rosario continúa acá, de brazos abiertos para todos los brasileños como si todos fuésemos hermanos.

El retorno a la Rosario fue tranquilo, pues como ya conocía la ciudad todo la fue más fácil. La sorpresa fue la mudanza del dirección de la universidad, pues el hotel que habíamos reservado estaba en las proximidades de la dirección anterior, lo que dificultó un poco. Cuanto el curso del doctorado, en no primero momento extrañé mucho la lenguaje, tuvo dificultad para comprender la lengua española. Yo tengo leídos muchos textos en español, ahora estoy comprendiendo mejor, aunque aún tenga dificultad, principalmente con la pronuncia. Gradativamente estoy mejorando y mi sintiendo más confortable en las clases del español. Supongo que hasta la defesa de la tese estará preparada.

Mi ciudade és muy diferente, pues Rosario tene una estructura bella. Rosario tene una cuantidade del hablantes á mais que mi ciudade, cá tene parcas belísimas, muy autorizada, las pesonas mayores, tene una cualidade del cuale muy bien. Cuando estoe mirando las pesonas majores, sacando mí madre y padre. Cuosta ma de morar aca, con mías padres.

Son cidades que tienen cosas en común. Sao Luis tiene un patrimonio arquitectónico mui antigo de colonización portuguesa. Rosario también tiene uno. Yo penso que el patrimonio es de colonización española Sao Luis tiene un clima fervoroso y Rosario templado.

Yo estoy en Rosario hay diez días. Ella es mucho bonita y tiene muchas plazas, arboleda, edificios y uno grande río conocido como Paraná. El tránsito es mucho ajetreado no obstante los conductores son mucho educados. Rosario tiene una universidad mucho respetado internacionalmente y también nacionalmente. Muchos brasileños estén acá. Yo particularmente, no mi adapté con la comida ventida en los parrillas. Yo soy grato al pueblo por la su hospitalidad.

Yo vivo en una ciudad histórica que se llama Sao Luis. Ella és tan rica culturalmente cuanto és Rosario en Argentina. Vine estudiar doctorado con mis amigas, estamos disfrutando mucho. Cuando hablé con mi madre, ella preguntó: \_¿Has salido mucho en Rosario? ¿La lengua es muy difícil de hablar? Yo respondí que no tengo dificultad y que estoy gostando mucho del Rosario, su gente es muy acogedora.

Ahora estoy de vuelta un paso más para cumplir otra etapa del mi doctorado. Yo ya sabía que en julio aquí faz mucho frío. Así es como conocí a Rosario, una ciudad tan fría en esta temporada. Fue también el Monumento de la bandera, lugar hermoso. Había grupos de varias ciudades que ofrecen sus danzas folclóricas. Me llevaré conmigo muchos recuerdos aquí.

En enero, cuando nos vinieron a ver el primer módulo del doctorado, el clima era muy similar a nuestra tierra, en torno a 32 grados, las personas se vestía de una manera única , donde cada uno tenía su propio estilo, el día estaba amaneciendo a las 6 y oscureciendo en torno a las 9. Ahora bien, este segundo módulo, los meses de julio, nosotras aterrizamos en aeropuerto de rosario tres horas de la mañana y el clima era muy frío en torno de tres grados. Ya las personas san vestida con ropa pesado para se agasalharem bien, se ve un desfile de moda, ahora el día esclarezca y el noitece igual, como ocurre todo el año en Brasil.

Alkmaar es una ciudad en Holanda, 40 km Nord (al norte) de Amsterdam y tiene 94.000 habitantes.

Pode llevar esta ciudad(se puede llegar a esta ciudad) por Schiphol, el Aeropuerto (aeropuerto) de Holanda, y despues en auto ó tren.

Alkmaar tiene unos hoteles (varios hoteles). Un hotel nuevo cerca del centro era en pasado palacio de justicia ( era en el pasado el Palacio de Justicia) y otro hotel nuevo era en pasado (en el pasado) una residencia de ancianos, 10 min caminar del centro (a 10' a pie desde el centro).

En verano el tiempo es mas lindo con temperaturas hasta 30 grados.

Visita unos de sus museos, como museo de queso, museo de cerveza, museo de calefacciones (calefacciones) o "Stedelijk Museum", museo de artes.

No se olvida mirar (no se olvide de mirar) el Ferio de Queso (la Feria del Queso). Este es los viernes en los meses abril hasta setiembre.

Alkmaar tiene algunos edificios antiguos. Per ejemplo "Grote Kerk" (Iglesia Grande) o Laurentiuskerk (Iglesia Lorenzo). Este edificio es de (del) Siglo 16. Este iglesia tiene dos organos. Uno es muy grande. Este organo esta en un placard gigante con postigos (los postigos) mas grandes del mudo. Estos postigos estan decorada de con figures de Biblia (están decorados con figuras de la Biblia). El otro organo es el mas viejo de Holanda, todavia puede tocar este organo. Este organo es de 1511, Con este (estos) dos órganos famosos, alkmaar es uno (una) de 9 (de las 9) ciudades de organo, juntos con Göteborg, Freiberg, Zaragoza, Roskilde, Toulouse, Lissabon, Innsbruck, y Treviso.

Otro edificio antiguo es el ayuntamiento, también de siglo 16. Una torre interesante es "De Waagtoren", el (la) torre en la plaza de mercado de queso. Escoucha a la hora en punta (Escucha la hora en punto) el carillon y mira las esculturas en la torre que mueven, se llaman "los cabelleros".

Alkmaar tiene muchos canales. Tome un barco para mirar la ciudad desde los canales.

Disfrute su centro con sus muchos negocios (negocios). Para la noche hay muchos restaurantes y bares. Alkmaar tiene un "barrio rojo", solamente una calle, Achterdam, pero para turistas interesante para mirar.

El mar es (está) muy cerca, solamente 7 km se encontra (encuentra) la costa con sus playas.

Para mi Alkmaar es mas linda en primavera y otoño. En primavera tiene posibilidades para mirar los campos de bulbos. Milliones (millones) de flores coloran (dan color al) el campo en todos los colores de arco iris y en un pueblo 8 km sur de esta ciudad hay mosaicos de jacintos.

En otoño, la ultima (el último) fin de semana de setiembre, es el Festival de Jazz. Y en octubre 8 la ciudad celebra "Alkmaar Ontzet" la victoria con la gente de alkmaar contra los españoles en 1573.

Pero en invierno tiene posibalidades para difrutar también. 30 km sur de Alkmaar hay una ciudad se llama Zaanstad y alla es "Zaanse Schans" un museo donde se demonstnan (demuestra) como haceros (se hacen los suecos) suecos y como se moleron (muele) el trigo en un molino.

30 km este de Alkmaar en Enkhuizen hay un museo de casas antiguos donde se demonstnan (muestra) como gente viveron en las costas del mar en pasado (como vivía la gente en la costas del mar hace muchos años).

Resumen de policial 22-09-06(Resumen Policial)

### El auto de Carrascal

Este artículo(artículo) era sobre un asesinato de una maestra.

Juan Pablo Carrascal confesó en detalles(detalle) como el maniató, violó(violó) y mató(mató a) Daniela Cintia Sparvoli a un camino rural(en un camino rural) en Cañada de Gomez.

Tambien conta(cuenta) este artículo que el asesinator(asesino) y victima hacian y la fecha del crimen. (También cuenta el artículo lo que hacían el asesino y la víctima y se da a conocer la fecha del crimen)

El trabajaba en una fabrica. El dueño de esta fabrica es el padre de su concubina.

El vivia con su concubina y los padres de su concubina,

Ella trabajaba como maestra. La maestra viajaba a su casa hacer dedo(haciendo dedo).

En 12 de mayo 2003(El 12 de Mayo de 2003) el vio a la y tenia idea para(él la vio y tuvo la idea de) tener sexo con ella.

Despues el crimen el robo su portafolio, rompió una photo(foto) y el documento de identidad y escondió debajo un armatorio(y lo escondió debajo de un armario).

La policía localicó a el asesinator(localizó al asesino) después estudiar(después de estudiar) muestras de ADN.

La policia deteneron el asesinator(detuvo al asesino) en Mendoza, 21 de septiembre 2006

### Victimas de trafico en Argentina.

Soy de Holanda y desde abril de 2003 vivo en Rosario, Argentina.

En los 3 y media años noté muchos diferentes en trafico con Holanda y Argentina.

Las autopistas per ejemplo son tranquilo, en Holanda son como cerca de Buenos Aires, con mucho trafico. Pero, pese al hecho de que hay menor trafico, pasa mas accidentes. Yo escuché un numero 7000 muertes por año, es mucho. Me pregunto, por que y como se puede evitar?

Yo sé unas soluciones. En la ciudad las automovilistas tiene que respetar los peatones y cyclistas. Ceder los peatones en paso de cebras y tener mas cuidado para cyclistas. Las automovilistas tiene que poner su cinterones, tambien su pasajeros delante y detras para su seguridad. Muchos victimas en motociclistas puede reducida si ponen su casco. Veo muchas sin.

Tecnico hay unos correcciones posible tambien. Cruces peligrosas transformar a rotondas. Ventajas son el trafico tiene que frenar (chance lesión en caso colisión mas pequeño) y semaforos no necesitan.

Afuera de ciudad ví unos accidentes tambien, mayoria unilateral. Creo son accidentes que tiene su causo en adormilarse de las automovilistas. En comparación que Holanda aca hay pocos lugares donde las automovilistas pueden descansar. Una solucion puede ser constuir mas sitios para aparcar acerca de servicios de estationes.

Tambien hay muchos accidentes frontál. Soluciones, pone sus luces de cruce, no adelantar en una curva, viaducto o si hay niebla y no manajar rapido.

Una otra solucion es transformar rutas con cruces y dos manos en autopistas y cruces a disniveles.

Espero los participanten de trafico y las autoridades unos de esta posibilidades realizan para reducir el numero de las victimas en trafico.

Christiaan Huygens (nació en La Haya el 14 de abril de 1629, murió el 8 de julio de 1695) era un astrónomo holandés asi como matemático y fisico.

La familia de Huygens era próspera y destacada. Su padre era diplomático y poeta.

Gracias a él, Christiaan tenia contactos con Descartes, un famoso matemático y filósofo.

Descartes tenia influencia en su educacion de matematica.

Huygens estudió derecho y matemática en la Universidad de Leiden desde 1645 hasta 1647, despues el continuó sus estudios en el Colegio de Naranja en Breda.

Porque su padre tambien tenia contactos con Mersenne, un otro matematico, el estuvo ocupado con la problema de una cuerda, colgando a sus extremos.

Desde entonces el tenia mas atencion de astronomia y la cálculo de la probabilidad.

Huygens y el pensionario de estatuder Johan de Witt contriburion un rol importante en la matematica de compañías de seguros.

Huygens espacialemente era famoso de sus teorias que luz existe en ondas, que estube en contradiccion con la teoria de Newton, que dice luz existe de particulas.

En 1655 descubrió Titan, la mas grande luna de Saturno

Tambien eximanaba los secciones de conica y hacia trabajo preparado para diferenciar y integrar.

La astronomia requiere la medicion exacta del tiempo por lo que dedicarse en este problema. El estudio el movimiento oscilatorio y en 1656 tenió su patente de primero reloj de péndulo.

Huygens tambien era el inventor de 31 tonos por oktavo en lugar de 12 medios tonos.

El el museo Teyler hay un organo basada de 31 tonos.

El museo Teyler es el museo mas viejo en Holanda, (1784) y se encontra en Haarlem.

Hola Daniela!

En este E-mail yo te quiero cantar (contar) mis vacaciones a (en) Mar del Plata.

Yo fue (fui) alla (allá) con mi esposa y mi mas joven hijastro (mi hijastro más joven).

Mi esposa reservó un departamento con 2 dormitorios y televisión (televisión) con cable por 8 dias.

Salimos en madrugada (la madrugada) de (del) 2 de febrero a las 3 para evitar retraso en Buenos Aires.

Buenos Aires (En Buenos Aires) es difícil (difícil) manejar en autopistas, porque no hay indicadores de lugares, pero solamente numeros de salidas.

Ely, mi esposa, encontré (encontró) la salida necesitamos (necesitábamos) y no tenemos (tuvimos) problemas para encontrar la salida correcta.

Llegamos en Mar del Plata (a Mar del Plata) a las once y cuarto, pero nuestro departamento fue ocupado.

Despues escuchamos que no tenemos (teníamos) habitacion, pero la dueña nos offró (ofreció) otro en otro edificio.

Este (esta) habitación fue (era) un poquito pequeño (pequeña) y tenía (tenía) televisión (televisión) sin cable (solamente 2 canales para mirar)

El tiempo durante nuestros vacaciones (nuestras vacaciones) no fue muy bien (bueno). Temperaturas bajas y viento. Ademas en algunos dias lluvia. Quando (Cuando) hay (había) sol fuimos en la playa (íbamos a la playa) en una carpa, durante mal tiempo fuimos a los restaurantes, mercados y una vez fuimos a las lomadas de Sierra de los padres.

En 10 (El 10) de febrero volvimos a Rosario. Una viaja (Un viaje) con muchos accidentes y embotellamientos.

Espero verte pronto en clase para hablar mis vacaciones (hablar de mis vacaciones).

CORPUS	TEXTO	OCURRENCIA VERBO		TIPO DESV_
INGLÉS	1	SOY	SER	EST
INGLÉS	1	ESTOY	ESTAR	EST
INGLÉS	1	ESTÁ	ESTAR	EST
INGLÉS	1	TIENEN	TENER	EST
INGLÉS	2	VOY	IR	EST
INGLÉS	3	SOY	SER	EST
INGLÉS	3	TENGO	TENER	EST
INGLÉS	3	TENGO	TENER	EST
INGLÉS	3	VAN	IR	EST
INGLÉS	3	QUIERÍAN	QUERER	DESV2
INGLÉS	4	QUIERO	QUERER	EST
INGLÉS	4	FUE	IR	EST
INGLÉS	4	ES	SER	EST
INGLÉS	4	VOY	IR	EST
INGLÉS	4	QUIERO	QUERER	EST
INGLÉS	4	PUEDES	PODER	EST
INGLÉS	5	TENGO	TENER	EST
INGLÉS	5	QUEREN	QUERER	DESV1
INGLÉS	5	QUIEREN	QUERER	EST
INGLÉS	5	TIENES	TENER	EST
INGLÉS	6	TIENNES	TENER	DESV3



INGLÉS	6	DARES	DAR	DESV3
INGLÉS	6	QUIERO	QUERER	EST
INGLÉS	6	QUIERES	QUERER	EST
INGLÉS	6	PUEDO	PODER	EST
INGLÉS	6	TENGO	TENER	EST
INGLÉS	6	VOY	IR	EST
INGLÉS	6	TENGO	TENER	EST
INGLÉS	6	SOY	SER	EST
INGLÉS	7	ESTARÁN	ESTAR	EST
INGLÉS	8	PUEDE	PODER	EST
INGLÉS	9	SOY	SER	EST
INGLÉS	9	HAY	SER	EST
INGLÉS	10	PODRÍA	PODER	EST
INGLÉS	10	SON	SER	EST
INGLÉS	10	SOY	SER	EST
INGLÉS	11	VOY	IR	EST
INGLÉS	12	TENGO	TENER	EST
INGLÉS	12	ESTÁN	ESTAR	EST
INGLÉS	13	SON	SER	EST
INGLÉS	13	QUIERO	QUERER	EST
INGLÉS	13	PREFIERO	PREFERIR	EST
INGLÉS	13	QUIERO	QUERER	EST
INGLÉS	14	ES	SER	EST
INGLÉS	15	QUIERES	QUERER	EST
INGLÉS	15	ES	SER	EST
INGLÉS	15	ES	SER	EST
INGLÉS	15	QUIERES	QUERER	EST

INGLÉS	15	PUEDES	PODER	EST
INGLÉS	16	QUIERO	QUERER	EST
INGLÉS	16	QUIERO	QUERER	EST
INGLÉS	17	ES	SER	EST
INGLÉS	18	VEIS	IR	DESV4
INGLÉS	19	QUIERO	QUERER	EST
INGLÉS	20	TENGO	TENER	EST
INGLÉS	20	VAN	IR	EST
INGLÉS	21	QUIERO	QUERER	EST
INGLÉS	21	PUEDES	PODER	EST
INGLÉS	22	SOY	SER	EST
INGLÉS	22	QUIERO	QUERER	EST
INGLÉS	22	ES	SER	EST
INGLÉS	22	SON	SER	EST
INGLÉS	22	QUEREMOS	QUERER	EST
INGLÉS	22	SOMOS	SER	EST
INGLÉS	23	HACIDO	HACER	DESV1
INGLÉS	23	HACIEMOS	HACER	DESV1
INGLÉS	23	HACIEMOS	HACER	DESV3
INGLÉS	23	HACÉS	HACER	EST
INGLÉS	23	QUIERO	QUERER	EST
INGLÉS	23	QUEREMOS	QUERER	EST
INGLÉS	23	TENEMOS	TENER	EST
INGLÉS	23	SOMOS	SER	EST
INGLÉS	23	PUEDE	PODER	EST
INGLÉS	23	SOY	SER	EST
INGLÉS	23	SOY	SER	EST

INGLÉS	23	QUIERO	QUERER	EST
INGLÉS	23	PONES	PONER	EST
INGLÉS	23	QUIERO	QUERER	EST
INGLÉS	23	HE VENIDO	VENIR	EST
INGLÉS	24	QUEREMOS	QUERER	EST
INGLÉS	24	ES	SER	EST
INGLÉS	24	TENEMOS	TENER	EST
INGLÉS	25	PUEDEMOS	PODER	DESV2
INGLÉS	25	QUIEREMOS	QUERER	DESV2
INGLÉS	25	SOMOS	SER	EST
INGLÉS	25	ES	SER	EST
INGLÉS	25	PUEDE	PODER	EST
INGLÉS	26	QUIERO	QUERER	EST
INGLÉS	26	QUIERO	QUERER	EST
INGLÉS	26	PENSIO	PENSAR	DESV1
INGLÉS	26	PENSIO	PENSAR	DESV3
INGLÉS	27	ES	SER	EST
INGLÉS	27	ES	SER	EST
INGLÉS	27	VOY	IR	EST
INGLÉS	27	SON	SER	EST
INGLÉS	27	QUERÍA	QUERER	EST
INGLÉS	27	SON	SER	EST
INGLÉS	27	QUIERO	QUERER	EST
INGLÉS	27	ES	SER	EST
INGLÉS	27	ES	SER	EST
INGLÉS	27	QUIERO	QUERER	EST
INGLÉS	27	SON	SER	EST

INGLÉS	27	SERÍA	SER	EST
INGLÉS	27	ES	SER	EST
INGLÉS	27	DORMIMOS	DORMIR	EST
INGLÉS	27	HAMOS VISITADO	HABER	DESV1
INGLÉS	27	HAMOS APRENDIDO	HABER	DESV1
INGLÉS	27	HE CONOSCADO	CONOCER	DESV3
INGLÉS	27	HE CONOSCADO	CONOCER	DESV4
INGLÉS	27	HA ESTIDO	ESTAR	DESV4
INGLÉS	28	DORMIMOS	DORMIR	EST
INGLÉS	28	HE CONOCIDO	CONOCER	EST
INGLÉS	28	FUI	IR	EST
INGLÉS	28	SON	SER	EST
INGLÉS	28	SON	SER	EST
INGLÉS	28	VAMOS	IR	EST
INGLÉS	28	PUEDA	PODER	EST
INGLÉS	28	ES	SER	EST
INGLÉS	28	QUIERO	QUERER	EST
INGLÉS	28	ES	SER	EST
INGLÉS	28	PUEDO	PODER	EST
INGLÉS	28	PUEDO	PODER	EST
INGLÉS	29	QUIERO	QUERER	EST
INGLÉS	29	SERÍA	SER	EST
INGLÉS	29	ES	SER	EST
INGLÉS	29	TENGO	TENER	EST
INGLÉS	29	SIENTO	SENTIR	EST
INGLÉS	29	ES	SER	EST

INGLÉS	29	DI	DAR	EST
INGLÉS	30	ESTOY	ESTAR	EST
INGLÉS	30	VOY	IR	EST
INGLÉS	30	HAGO	HACER	EST
INGLÉS	30	ESTÁ	ESTAR	EST
INGLÉS	30	ES	SER	EST
INGLÉS	30	QUIERO	QUERER	EST
INGLÉS	30	FUEMOS	IR	DESV3

CORPUS	TEXTO	OCURRENCIA VERBO		TIPO DESV_
ALEMÁN	1	SUGERÍÓ	SUGERIR	DESV1
ALEMÁN	1	RUEGUÍA	ROGAR	DESV2
ALEMÁN	1	RUEGUÍA	ROGAR	DESV4
ALEMÁN	1	TRAYERA	TRAER	DESV3
ALEMÁN	1	TRAYESE	TRAER	DESV3
ALEMÁN	1	PONDIERA	PONER	DESV3
ALEMÁN	1	PAGUERAN	PAGAR	DESV3
ALEMÁN	1	PAGUERAN	PAGAR	DESV4
ALEMÁN	1	PIDIERA	PEDIR	EST
ALEMÁN	1	DIJO	DECIR	EST
ALEMÁN	1	DIJE	DECIR	EST
ALEMÁN	1	QUERÍA	QUERER	EST
ALEMÁN	1	DIJO	DECIR	EST
ALEMÁN	1	PAGABAN	PAGAR	EST

ALEMÁN	1	DIJO	DECIR	EST
ALEMÁN	1	DIJE	DECIR	EST
ALEMÁN	1	PODÍAN	PODER	EST
ALEMÁN	1	DIJE	DECIR	EST
ALEMÁN	1	DIJE	DECIR	EST
ALEMÁN	1	TUVIERAN	TENER	EST
ALEMÁN	2	RECOMIENDA	RECOMENDAR	EST
ALEMÁN	2	SUGIERE	SUGERIR	EST
ALEMÁN	3	SABEN	SABER	EST
ALEMÁN	3	ESTOY	ESTAR	EST
ALEMÁN	3	TENGO	TENER	EST
ALEMÁN	3	QUIERO	QUERER	EST
ALEMÁN	3	DIJE	DECIR	EST
ALEMÁN	3	HAYA QUERIDO	QUERER	EST
ALEMÁN	3	DIJERON	DECIR	EST
ALEMÁN	3	RECOMENDARON	RECOMENDAR	EST
ALEMÁN	3	ELIGIERA	ELEGIR	EST
ALEMÁN	3	SÉ	SABER	EST
ALEMÁN	3	PROPONÍ	PROPONER	DESV1
ALEMÁN	3	SACÉ	SACAR	DESV1
ALEMÁN	3	AGARÍA	HACER	DESV3
ALEMÁN	4	TENÍA	TENER	EST
ALEMÁN	4	EMPEZAMOS	EMPEZAR	EST
ALEMÁN	4	TIENEN	TENER	EST
ALEMÁN	4	EMPEZÓ	EMPEZAR	EST
ALEMÁN	4	TRAJERON	TRAER	EST
ALEMÁN	4	ERA	SER	EST

<b>ALEMÁN</b>	4	DORMÍ	DORMIR	EST
<b>ALEMÁN</b>	4	HABÍA VISTO	VER	EST
<b>ALEMÁN</b>	4	FUIMOS	IR	EST
<b>ALEMÁN</b>	4	FUIMOS	IR	EST
<b>ALEMÁN</b>	4	FUE	IR	EST
<b>ALEMÁN</b>	4	PUDE	PODER	EST
<b>ALEMÁN</b>	4	SENTÍ	SENTIR	EST
<b>ALEMÁN</b>	4	SENTÉ	SENTARSE	EST
<b>ALEMÁN</b>	4	ESTUBEN	ESTAR	DESV4
<b>ALEMÁN</b>	4	QUERÍA	QUERER	EST
<b>ALEMÁN</b>	4	HAY	HABER	EST
<b>ALEMÁN</b>	5	SACAMOS	SACAR	EST
<b>ALEMÁN</b>	6	FUERON	IR	EST
<b>ALEMÁN</b>	6	TUVE	TENER	EST
<b>ALEMÁN</b>	6	QUERABA	QUERER	DESV4
<b>ALEMÁN</b>	6	PEDO	PEDIR	DESV1
<b>ALEMÁN</b>	6	ESTOY	ESTAR	EST
<b>ALEMÁN</b>	7	FUE	IR	EST
<b>ALEMÁN</b>	7	DIO	DAR	EST
<b>ALEMÁN</b>	8	PIDERON	PEDIR	DESV1
<b>ALEMÁN</b>	8	TRAERON	TRAER	DESV1
<b>ALEMÁN</b>	8	VENIERON	VENIR	DESV1
<b>ALEMÁN</b>	8	HACISTE	HACER	DESV1
<b>ALEMÁN</b>	9	PREFERIO	PREFERIR	DESV1
<b>ALEMÁN</b>	9	PREFERIO	PREFERIR	DESV4
<b>ALEMÁN</b>	10	VOLVE	VOLVER	DESV1
<b>ALEMÁN</b>	10	JUGA	JUGAR	DESV1

ALEMÁN	11	TENÍA	TENER	EST
ALEMÁN	11	EMPEZÓ	EMPEZAR	EST
ALEMÁN	11	EMPEZÓ	EMPEZAR	EST
ALEMÁN	11	SABÍA	SABER	EST
ALEMÁN	11	DIO	DAR	EST
ALEMÁN	11	SENTÍÓ	SENTIR	DESV4
ALEMÁN	12	PUSQUE	BUSCAR	DESV3
ALEMÁN	12	EQUENTRE	ENCONTRAR	DESV3
ALEMÁN	12	TENÉS	TENER	EST
ALEMÁN	12	TIENE	TENER	EST
ALEMÁN	13	DOLO	DOLER	DESV4
ALEMÁN	13	ERA	SER	EST
ALEMÁN	13	HACE	HACER	EST
ALEMÁN	13	DIJO	DECIR	EST
ALEMÁN	13	DIJO	DECIR	EST
ALEMÁN	13	SABRÍAN	SABER	EST
ALEMÁN	13	TENÍA	TENER	EST
ALEMÁN	13	TENÍA	TENER	EST
ALEMÁN	13	SEA	SER	EST
ALEMÁN	13	ESTABA	ESTAR	EST
ALEMÁN	13	TUVO	TENER	EST
ALEMÁN	13	ERAN	SER	EST
ALEMÁN	13	ERA	SER	EST
ALEMÁN	13	HUBO	HABER	EST
ALEMÁN	13	HABÍA	HABER	EST
ALEMÁN	13	TIENEN	TENER	EST
ALEMÁN	13	TIENEN	TENER	EST



ALEMÁN	13	SON	SER	EST
ALEMÁN	13	SON	SER	EST
ALEMÁN	13	ERA	SER	EST
ALEMÁN	13	ES	SER	EST
ALEMÁN	14	HAY	HABER	EST
ALEMÁN	14	ES	SER	EST
ALEMÁN	14	HABÍA CONOCIDO	CONOCER	EST
ALEMÁN	14	ESTÁ	ESTAR	EST
ALEMÁN	14	CONOCÍA	CONOCER	EST
ALEMÁN	14	ESTUVO	ESTAR	EST
ALEMÁN	14	ESTUVO	ESTAR	EST
ALEMÁN	14	FUE	IR	EST
ALEMÁN	14	VOLVIMOS	VOLVER	EST
ALEMÁN	15	HABÍAN	HABER	DESV4
ALEMÁN	15	FUE	SER	EST
ALEMÁN	15	DIJE	DECIR	EST
ALEMÁN	15	SERÍA	SER	EST
ALEMÁN	15	ESTABA	ESTAR	EST
ALEMÁN	15	ESTABA	ESTAR	EST
ALEMÁN	15	PODÍA	PODER	EST
ALEMÁN	15	PODÍAMOS	PODER	EST
ALEMÁN	15	QUERÍAMOS	QUERER	EST
ALEMÁN	15	HAY	HABER	EST
ALEMÁN	15	FUERON	IR	EST
ALEMÁN	15	FUE	SER	EST
ALEMÁN	15	TIENE	TENER	EST
ALEMÁN	15	ROMPIERON	ROMPER	EST

ALEMÁN	15	HACÍA	HACER	EST
ALEMÁN	15	HACÍA	HACER	EST
ALEMÁN	15	DIMOS	DAR	EST
ALEMÁN	15	TENÍAMOS	TENER	EST
ALEMÁN	15	HICIMOS	HACER	EST
ALEMÁN	15	ES	SER	EST
ALEMÁN	15	SENTÍA	SENTIR	EST
ALEMÁN	15	ESTABA	ESTAR	EST
ALEMÁN	15	ESTÁBAMOS	ESTAR	EST
ALEMÁN	15	PARECÍAMOS	PARECER	EST
ALEMÁN	15	HABÍAMOS TRAÍDO	TRAER	EST
ALEMÁN	15	FUIMOS	IR	EST
ALEMÁN	16	ERA	SER	EST
ALEMÁN	16	SABE	SABER	EST
ALEMÁN	16	ERA	SER	EST
ALEMÁN	16	FUERON	IR	EST
ALEMÁN	16	TENÍAN	TENER	EST
ALEMÁN	16	FUIMOS	IR	EST
ALEMÁN	16	TENEMOS	TENER	EST
ALEMÁN	17	FUE	SER	EST
ALEMÁN	17	ENCONTRÉ	ENCONTRAR	EST
ALEMÁN	17	CONOCÍ	CONOCER	EST
ALEMÁN	17	RECORDAMOS	RECORDAR	EST
ALEMÁN	17	CONOCÍAMOS	CONOCER	EST
ALEMÁN	17	FUE	ERA	EST
ALEMÁN	17	CONOCÍ	CONOCER	EST

<b>ALEMÁN</b>	18	ERA	SER	EST
<b>ALEMÁN</b>	18	ERA	SER	EST
<b>ALEMÁN</b>	18	VINE	VENIR	EST
<b>ALEMÁN</b>	18	HABÍA	HABER	EST
<b>ALEMÁN</b>	18	SENTÍ	SENTIR	EST
<b>ALEMÁN</b>	18	HACÍA	HACER	EST
<b>ALEMÁN</b>	18	HABÍA SALIDO	SALIR	EST
<b>ALEMÁN</b>	18	ERA	SER	EST
<b>ALEMÁN</b>	18	HABÍA SABIDO	SABER	EST
<b>ALEMÁN</b>	18	SABÍA	SABER	EST
<b>ALEMÁN</b>	18	HABÍA	HABER	EST
<b>ALEMÁN</b>	18	FUE	SER	EST
<b>ALEMÁN</b>	18	ERA	SER	EST
<b>ALEMÁN</b>	19	HACE	HACER	EST
<b>ALEMÁN</b>	19	FUI	IR	EST
<b>ALEMÁN</b>	19	TENÍA	TENER	EST
<b>ALEMÁN</b>	19	FUIMOS	IR	EST
<b>ALEMÁN</b>	19	HIZO	HACER	EST
<b>ALEMÁN</b>	19	TUVIMOS	TENER	EST
<b>ALEMÁN</b>	19	PUSIMOS	PONER	EST
<b>ALEMÁN</b>	19	ESTUVIMOS	ESTAR	EST
<b>ALEMÁN</b>	20	ESTUVO	ESTAR	EST
<b>ALEMÁN</b>	20	TENÍA	TENER	EST
<b>ALEMÁN</b>	20	HABÍA VISTO	VER	EST
<b>ALEMÁN</b>	20	PUDE	PODER	EST
<b>ALEMÁN</b>	20	CAREZCO	CARECER	EST
<b>ALEMÁN</b>	20	FUI	IR	EST

<b>ALEMÁN</b>	20	TENÍA	TENER	EST
<b>ALEMÁN</b>	20	DI	DAR	EST
<b>ALEMÁN</b>	20	ANDUVE	ANDAR	EST
<b>ALEMÁN</b>	20	TENÍA	TENER	EST
<b>ALEMÁN</b>	20	HACÍA	HACER	EST
<b>ALEMÁN</b>	21	RECUERDA	RECORDAR	EST
<b>ALEMÁN</b>	21	TENÍAMOS	TENER	EST
<b>ALEMÁN</b>	21	PUEDE	PODER	EST
<b>ALEMÁN</b>	21	ES	SER	EST
<b>ALEMÁN</b>	21	VIENE	VENIR	EST
<b>ALEMÁN</b>	21	DIRÍA	DECIR	EST
<b>ALEMÁN</b>	21	SENTÍ	SENTIR	EST
<b>ALEMÁN</b>	21	FUE	IR	EST
<b>ALEMÁN</b>	21	HAYAN SIDO	SER	EST
<b>ALEMÁN</b>	21	ENCUENTRAN	ENCONTRAR	EST
<b>ALEMÁN</b>	21	SUELEN	SOLER	EST
<b>ALEMÁN</b>	21	ES	SER	EST
<b>ALEMÁN</b>	21	EMPIEZAN	EMPEZAR	EST
<b>ALEMÁN</b>	21	SABE	SABER	EST
<b>ALEMÁN</b>	21	ES	SER	EST
<b>ALEMÁN</b>	21	SIENTE	SENTIR	EST
<b>ALEMÁN</b>	21	VA	IR	EST

CORPUS	TEXTO	OCURRENCIA VERBO		TIPO DESV_
HOLANDÉS	1	ES	SER	EST
HOLANDÉS	1	TIENE	TENER	EST
HOLANDÉS	1	TIENE	TENER	EST
HOLANDÉS	1	ERA	SER	EST
HOLANDÉS	1	ERA	SER	EST
HOLANDÉS	1	ES	SER	EST
HOLANDÉS	1	ES	SER	EST
HOLANDÉS	1	TIENE	TENER	EST
HOLANDÉS	1	ES	SER	EST
HOLANDÉS	1	TIENE	TENER	EST
HOLANDÉS	1	ES	SER	EST
HOLANDÉS	1	ESTÁ	ESTAR	EST
HOLANDÉS	1	ESTÁN	ESTAR	EST
HOLANDÉS	1	ES	SER	EST
HOLANDÉS	1	PUEDE	PODER	EST
HOLANDÉS	1	ES	SER	EST
HOLANDÉS	1	ES	SER	EST
HOLANDÉS	1	MUEVEN	MOVER	EST
HOLANDÉS	1	TIENE	TENER	EST
HOLANDÉS	1	HAY	HABER	EST
HOLANDÉS	1	TIENE	TENER	EST
HOLANDÉS	1	ES	SER	EST
HOLANDÉS	1	ES	SER	EST
HOLANDÉS	1	TIENE	TENER	EST
HOLANDÉS	1	HAY	HABER	EST
HOLANDÉS	1	ES	SER	EST

HOLANDÉS	1	TIENE	TENER	EST
HOLANDÉS	1	HAY	HABER	EST
HOLANDÉS	1	HAY	HABER	EST
HOLANDÉS	1	PODE	PODER	DESV1
HOLANDÉS	1	ENCONTRA	ENCONTRAR	DESV1
HOLANDÉS	1	HACEROS	HACER	DESV1
HOLANDÉS	1	MOLERON	MOLER	DESV1
HOLANDÉS	1	DEMONSTRAN	MOSTRAR	DESV3
HOLANDÉS	1	DEMONSTRAN	DEMOSTRAR	DESV3
HOLANDÉS	2	ERA	SER	EST
HOLANDÉS	2	HACÍAN	HACER	EST
HOLANDÉS	2	ES	SER	EST
HOLANDÉS	2	TENÍA	TENER	EST
HOLANDÉS	2	ES	SER	EST
HOLANDÉS	2	CONTA	CONTAR	DESV1
HOLANDÉS	2	DETENERON	DETENER	DESV1
HOLANDÉS	2	DETENERON	DETENER	DESV3
HOLANDÉS	2	LOCALICÓ	LOCALIZAR	DESV3
HOLANDÉS	3	SOY	SER	EST
HOLANDÉS	3	SON	SER	EST
HOLANDÉS	3	HAY	HABER	EST
HOLANDÉS	3	ES	SER	EST
HOLANDÉS	3	PUEDE	PODER	EST
HOLANDÉS	3	SÉ	SABER	EST
HOLANDÉS	3	PUEDE	PODER	EST
HOLANDÉS	3	HAY	HABER	EST
HOLANDÉS	3	TIENE	TENER	EST

HOLANDÉS	3	SON	SER	EST
HOLANDÉS	3	HAY	HABER	EST
HOLANDÉS	3	PUEDEN	PODER	EST
HOLANDÉS	3	PUEDE	PODER	EST
HOLANDÉS	3	HAY	HABER	EST
HOLANDÉS	3	HAY	HABER	EST
HOLANDÉS	3	ES	SER	EST
HOLANDÉS	3	NACÍÓ	NACER	EST
HOLANDÉS	4	MURIÓ	MORIR	EST
HOLANDÉS	4	ERA	SER	EST
HOLANDÉS	4	ERA	SER	EST
HOLANDÉS	4	TENÍA	TENER	EST
HOLANDÉS	4	TENÍA	TENER	EST
HOLANDÉS	4	TENÍA	TENER	EST
HOLANDÉS	4	ESTUVO	ESTAR	EST
HOLANDÉS	4	TENÍA	TENER	EST
HOLANDÉS	4	ERA	SER	EST
HOLANDÉS	4	HACÍA	HACER	EST
HOLANDÉS	4	ERA	SER	EST
HOLANDÉS	4	HAY	HABER	EST
HOLANDÉS	4	ES	SER	EST
HOLANDÉS	4	TENÍÓ	TENER	EST
HOLANDÉS	4	ENCONTRA	ENCONTRAR	DESV1
HOLANDÉS	4	CONTRIBURION	CONTRIBUIR	DESV1
HOLANDÉS	4	REQUIRE	REQUERIR	DESV3
HOLANDÉS	4	QUIERO	QUERER	DESV3
HOLANDÉS	5	ES	SER	EST

HOLANDES	5	HAY	HABER	EST
HOLANDES	5	FUE	SER	EST
HOLANDES	5	HAY	HABER	EST
HOLANDES	5	FUIMOS	IR	EST
HOLANDES	5	FUIMOS	IR	EST
HOLANDES	5	FUIMOS	IR	EST
HOLANDES	5	VOLVIMOS	VOLVER	EST
HOLANDES	5	TENIMOS	TENER	EST
HOLANDES	5	TENIMOS	TENER	DESV1
HOLANDES	5	TENIMOS	TENER	DESV1
HOLANDES	5	TENIÓ	TENER	DESV1
HOLANDES	5	OFFRÓ	OFRECER	DESV1
HOLANDES	5	TENIÓ	TENER	DESV3
HOLANDES	5	TENIÓ	TENER	DESV4
HOLANDES	6	VOY	IR	DESV1
HOLANDES	7	PIENSO	PENSAR	EST
HOLANDES	7	TENEMOS	TENER	EST
HOLANDES	7	TENÍS	TENER	DESV4

CORPUS	TEXTO	OCURRENCIA VERBO		TIPO DESV_
FRANCÉS	1	ESTUVE	ESTAR	EST
FRANCÉS	1	ESTOY	ESTAR	EST
FRANCÉS	1	TUVE	TENER	EST
FRANCÉS	1	SERÍA	SER	EST
FRANCÉS	1	QUIERO	QUERER	EST



FRANCÉS	2	ESTOY	ESTAR	EST
FRANCÉS	3	ERA	SER	EST
FRANCÉS	3	ES	SER	EST
FRANCÉS	3	TIENE	TENER	EST
FRANCÉS	3	TIENEN	TENER	EST
FRANCÉS	3	SEA	SER	EST
FRANCÉS	4	QUIERE	QUERER	EST
FRANCÉS	4	QUIERE	QUERER	EST
FRANCÉS	4	HACÍA	HACER	EST
FRANCÉS	4	VEÍAN	VER	EST
FRANCÉS	4	Eran	SER	EST
FRANCÉS	4	Eran	SER	EST
FRANCÉS	5	QUERÍAN	QUERER	EST
FRANCÉS	5	DIJO	DECIR	EST
FRANCÉS	5	FUERON	IR	EST
FRANCÉS	6	Eran	SER	EST
FRANCÉS	6	Eran	SER	EST
FRANCÉS	6	HACEN	HACER	EST
FRANCÉS	6	VEN	VER	EST
FRANCÉS	7	QUIERES	QUERER	EST
FRANCÉS	7	HAY	HABER	EST
FRANCÉS	7	VA	IR	EST
FRANCÉS	8	NACÍÓ	NACER	EST
FRANCÉS	8	TENÍA	TENER	EST
FRANCÉS	8	TENÍAN	TENER	EST
FRANCÉS	8	HIZO	HACER	EST
FRANCÉS	9	DERRITA	DERRETIR	EST

FRANCÉS	9	PONGA	PONER	EST
FRANCÉS	9	PONGA	PONER	EST
FRANCÉS	9	COLOCE	COLOCAR	DESV1
FRANCÉS	9	PICE	PICAR	DESV1
FRANCÉS	9	CALENTE	CALENTAR	DESV1
FRANCÉS	9	PIQE	PICAR	DESV3
FRANCÉS	9	FREE	FREIR	DESV1
FRANCÉS	9	FREE	FREIR	DESV4
FRANCÉS	9	AGREGE	AGREGAR	DESV3
FRANCÉS	9	COLOQE	COLOCAR	DESV3
FRANCÉS	10	ESTÁS	ESTAR	EST
FRANCÉS	10	ES	SER	EST
FRANCÉS	11	SACE	SACAR	DESV1
FRANCÉS	12	PUEDE	PODER	EST
FRANCÉS	12	PUEDO	PODER	EST
FRANCÉS	12	ESTÁN	ESTAR	EST
FRANCÉS	12	ESTOY	ESTAR	EST
FRANCÉS	13	FUERON	IR	EST
FRANCÉS	13	FUERON	IR	EST
FRANCÉS	14	ES	SER	EST
FRANCÉS	14	TENGO	TENER	EST
FRANCÉS	14	TIENE	TENER	EST
FRANCÉS	15	FUERON	IR	EST
FRANCÉS	15	SACIERON	SACAR	DESV4
FRANCÉS	15	CONCÍ	CONOCER	DESV3
FRANCÉS	15	FUERON	IR	EST
FRANCÉS	15	CONOCIERON	CONOCER	EST

FRANCÉS	15	FUI	IR	EST
FRANCÉS	15	FUI	IR	EST
FRANCÉS	16	FUI	IR	EST
FRANCÉS	16	ESTOY	ESTAR	EST
FRANCÉS	17	QUISO	QUERER	EST
FRANCÉS	17	FUE	IR	EST
FRANCÉS	17	VINO	VENIR	EST
FRANCÉS	18	PARECE	PARECER	EST
FRANCÉS	18	ES	SER	EST
FRANCÉS	18	NACÍÓ	NACER	EST
FRANCÉS	18	PUDIERON	PODER	EST
FRANCÉS	18	ERA	SER	EST
FRANCÉS	18	SON	SER	EST
FRANCÉS	18	VINO	VENIR	EST
FRANCÉS	19	TIENE	TENER	EST
FRANCÉS	19	ES	SER	EST
FRANCÉS	19	ES	SER	EST
FRANCÉS	19	TIENE	TENER	EST
FRANCÉS	20	HAY	HABER	EST
FRANCÉS	20	SON	SER	EST
FRANCÉS	20	HAY	HABER	EST
FRANCÉS	20	ES	SER	EST
FRANCÉS	20	TENÉ	TENER	DESV4
FRANCÉS	21	ESTÁ	ESTAR	EST
FRANCÉS	21	HAY	HABER	EST
FRANCÉS	21	QUIERE	QUERER	EST
FRANCÉS	21	QUIERE	QUERER	EST

FRANCÉS	21	ES	SER	EST
FRANCÉS	22	PONÉ	PONER	EST
FRANCÉS	23	TIENE	TENER	EST
FRANCÉS	23	TIENE	TENER	EST
FRANCÉS	23	ES	SER	EST
FRANCÉS	23	ES	SER	EST
FRANCÉS	23	HAY	HABER	EST
FRANCÉS	24	QUISO	QUERER	EST
FRANCÉS	24	TUVO	TENER	EST
FRANCÉS	24	QUERÍAN	QUERER	EST
FRANCÉS	24	DIERON	DAR	EST
FRANCÉS	24	ENCONTRÓ	ENCONTRAR	EST
FRANCÉS	24	VIO	VER	EST
FRANCÉS	24	QUISIERON	QUERER	EST
FRANCÉS	24	TENÍAN	TENER	EST
FRANCÉS	24	ESTABA	ESTAR	EST
FRANCÉS	25	QUISE	QUERER	EST
FRANCÉS	25	ERA	SER	EST
FRANCÉS	25	ESTABA	ESTAR	EST
FRANCÉS	25	HACÍA	HACER	EST
FRANCÉS	25	ESTABA	ESTAR	EST
FRANCÉS	25	IBA	IR	EST
FRANCÉS	25	HABÍA	HABER	EST
FRANCÉS	25	TENÍA	TENER	EST
FRANCÉS	26	ESTUVA	ESTAR	DESV4
FRANCÉS	26	VIA	VER	DESV4
FRANCÉS	27	FUE	IR	EST

FRANCÉS	27	DIRIGO	DIRIGIR	DESV3
FRANCÉS	27	VIO	VER	EST
FRANCÉS	28	QUIERO	QUERER	EST
FRANCÉS	28	QUIERO	QUERER	EST
FRANCÉS	28	ESTÁ	ESTAR	EST
FRANCÉS	28	TIENE	TENER	EST
FRANCÉS	29	QUIERE	QUERER	EST
FRANCÉS	29	ES	SER	EST
FRANCÉS	29	DESCUBREN	DESCUBRIR	EST
FRANCÉS	29	PUEDE	PODER	EST
FRANCÉS	29	ES	SER	EST
FRANCÉS	29	QUIERE	QUERER	EST
FRANCÉS	29	QUIERE	QUERER	EST
FRANCÉS	29	ES	SER	EST
FRANCÉS	30	VUELVÉ	VOLVER	DESV2
FRANCÉS	30	VUELVÉ	VOLVER	DESV4
FRANCÉS	29	EMPEZA	EMPEZAR	DESV1
FRANCÉS	29	EMPEZO	EMPEZAR	DESV2
FRANCÉS	30	SON	SER	EST
FRANCÉS	30	ESTÁ	ESTAR	EST
FRANCÉS	30	TENGO	TENER	EST
FRANCÉS	30	PONGO	PONER	EST
FRANCÉS	30	VOY	IR	EST
FRANCÉS	30	TENGO	TENER	EST
FRANCÉS	30	HAY	HABER	EST
FRANCÉS	30	ESTÁ	ESTAR	EST
FRANCÉS	30	ENCONTRÉ	ENCONTRAR	EST

FRANCÉS	31	ES	SER	EST
FRANCÉS	31	TIENE	TENER	EST
FRANCÉS	31	PIDE	PEDIR	EST
FRANCÉS	31	CONDUZCA	CONducIR	EST
FRANCÉS	31	FUE	IR	EST
FRANCÉS	31	VIO	VER	EST
FRANCÉS	31	PODEMOS	PODER	EST
FRANCÉS	31	ES	SER	EST
FRANCÉS	31	TENEMOS	TENER	EST
FRANCÉS	31	MUESTRA	MoSTRAR	EST
FRANCÉS	31	PIENSO	PENSAR	EST
FRANCÉS	31	SON	SER	EST
FRANCÉS	31	PUEDEN	PODER	EST
FRANCÉS	32	TENGO	TENER	EST
FRANCÉS	32	TENGO	TENER	EST
FRANCÉS	32	ESTÁN	ESTAR	EST
FRANCÉS	32	ES	SER	EST
FRANCÉS	32	ES	SER	EST
FRANCÉS	32	ES	SER	EST
FRANCÉS	32	ENCUENTRAN	ENCONTRAR	EST
FRANCÉS	32	SUPONEN	SUPONER	EST
FRANCÉS	32	QUIEREN	QUERER	EST
FRANCÉS	32	PODEMOS	PODER	EST
FRANCÉS	32	CONSTRUYEN	CONSTRUIR	EST
FRANCÉS	32	HAY	HABER	EST
FRANCÉS	32	ES	SER	EST
FRANCÉS	32	PARECE	PARECER	EST

FRANCÉS	32	FUE	IR	EST
FRANCÉS	32	PODEMOS	PODER	EST
FRANCÉS	32	PODRÁN	PODER	EST
FRANCÉS	32	PUEDE	PODER	EST
FRANCÉS	33	AMENAZA	AMENAZAR	EST
FRANCÉS	33	RENOVAN	RENOVAR	DESV1
FRANCÉS	34	FUE	SER	EST
FRANCÉS	34	ES	SER	EST
FRANCÉS	34	SEA	SER	EST
FRANCÉS	34	PARECE	PARECER	EST
FRANCÉS	34	FUE	SER	EST
FRANCÉS	34	QUIEREN	QUERER	EST
FRANCÉS	34	PODEMOS	PODER	EST
FRANCÉS	34	ES	SER	EST
FRANCÉS	34	TENÍA	TENER	EST
FRANCÉS	34	ES	SER	EST
FRANCÉS	35	TENGO	TENER	EST
FRANCÉS	36	PARECE	PARECER	EST
FRANCÉS	36	ES	SER	EST
FRANCÉS	36	QUIERO	QUERER	EST
FRANCÉS	36	FUE	SER	EST
FRANCÉS	36	TENGO	TENER	EST
FRANCÉS	37	ES	SER	EST
FRANCÉS	37	HACE	HACER	EST
FRANCÉS	37	VOY	IR	EST
FRANCÉS	37	ES	SER	EST
FRANCÉS	37	ES	SER	EST

FRANCÉS	37	SON	SER	EST
FRANCÉS	37	ERA	SER	EST
FRANCÉS	38	SOY	SER	EST
FRANCÉS	38	ES	SER	EST
FRANCÉS	38	SON	SER	EST
FRANCÉS	38	COMENCO	COMENZAR	DESV1
FRANCÉS	39	SOY	SER	EST
FRANCÉS	39	TIENE	TENER	EST
FRANCÉS	39	SON	SER	EST
FRANCÉS	39	QUIERO	QUERER	EST
FRANCÉS	40	ESTOY	ESTAR	EST
FRANCÉS	40	FUE	ES	EST
FRANCÉS	40	ESTOY	ESTAR	EST
FRANCÉS	40	VAS	IR	EST
FRANCÉS	40	HACE	HACER	EST
FRANCÉS	40	ES	SER	EST
FRANCÉS	40	PUEDO	PODER	EST
FRANCÉS	40	ESTÁ	ESTAR	EST
FRANCÉS	40	ESTÁ	ESTAR	EST
FRANCÉS	40	ESTÁ	ESTAR	EST
FRANCÉS	41	ESTÁS	ESTAR	EST
FRANCÉS	41	ESTOY	ESTAR	EST
FRANCÉS	41	ES	SER	EST
FRANCÉS	41	PUEDO	PODER	EST
FRANCÉS	41	HACÉS	HACER	EST
FRANCÉS	41	QUIERO	QUERER	EST
FRANCÉS	41	ESTÁ	ESTAR	EST



FRANCÉS	42	ES	SER	EST
FRANCÉS	42	ENCUENTRE	ENCONTRAR	EST
FRANCÉS	42	EMPIEZA	EMPEZAR	EST
FRANCÉS	42	TENDRÍAN	TENER	EST
FRANCÉS	42	DIJISTE	DECIR	EST
FRANCÉS	42	QUERÍAS	QUERER	EST
FRANCÉS	42	ES	SER	EST
FRANCÉS	42	PODÉS	PODER	EST
FRANCÉS	42	ES	SER	EST
FRANCÉS	42	QUIERES	QUERER	EST
FRANCÉS	42	ENCONTREMOS	ENCONTRAR	EST
FRANCÉS	43	SOMOS	SER	EST
FRANCÉS	43	ESTARÁN	ESTAR	EST
FRANCÉS	43	ES	SER	EST
FRANCÉS	43	HAY	HABER	EST
FRANCÉS	43	SON	SER	EST
FRANCÉS	43	SOMOS	SER	EST
FRANCÉS	43	TENGES	TENER	DESV4
FRANCÉS	44	SOY	SER	EST
FRANCÉS	44	FUI	IR	EST
FRANCÉS	44	QUIERO	QUERER	EST
FRANCÉS	44	QUIERO	QUERER	EST
FRANCÉS	44	TENGO	TENER	EST
FRANCÉS	44	PUEDÁIS	PODER	DESV4
FRANCÉS	45	TENGO	TENER	EST
FRANCÉS	45	QUIERO	QUERER	EST
FRANCÉS	45	HAY	HABER	EST

<b>FRANCÉS</b>	45	PUEDO	PODER	EST
<b>FRANCÉS</b>	46	QUIERO	QUERER	EST
<b>FRANCÉS</b>	46	ES	SER	EST
<b>FRANCÉS</b>	46	ES	SER	EST
<b>FRANCÉS</b>	46	RECOMIENDO	RECOMENDAR	EST
<b>FRANCÉS</b>	46	ES	SER	EST
<b>FRANCÉS</b>	46	SON	SER	EST
<b>FRANCÉS</b>	47	ES	SER	EST
<b>FRANCÉS</b>	47	SON	SER	EST
<b>FRANCÉS</b>	47	QUIERO	QUERER	EST
<b>FRANCÉS</b>	48	FUI	IR	EST
<b>FRANCÉS</b>	48	ES	SER	EST
<b>FRANCÉS</b>	48	ESTOY	ESTAR	EST
<b>FRANCÉS</b>	48	VAN	IR	EST
<b>FRANCÉS</b>	48	QUIERO	QUERER	EST
<b>FRANCÉS</b>	48	PUEDEN	PODER	EST

<b>CORPUS</b>	<b>TEXTO</b>	<b>OCURRENCIA</b>	<b>VERBO</b>	<b>TIPO DESV-</b>
<b>PORTUGUÉS</b>	1	TIENE	TENER	EST
<b>PORTUGUÉS</b>	1	PROPOSTO	PROPONER	DESV1
<b>PORTUGUÉS</b>	1	EXPONEMOS	EXPONER	EST
<b>PORTUGUÉS</b>	2	INCRUYO	INCLUIR	DESV3
<b>PORTUGUÉS</b>	3	ESTÁ	ESTAR	EST
<b>PORTUGUÉS</b>	3	ES	SER	EST

PORTUGUÉS	3	ESTÁ	ESTAR	EST
PORTUGUÉS	3	QUIERO	QUERER	EST
PORTUGUÉS	3	SALGA	SALIR	EST
PORTUGUÉS	4	ARREPINTINRON	ARREPENTIRSE	DESV3
PORTUGUÉS	4	ARREPENTIMOS	ARREPENTIRSE	EST
PORTUGUÉS	5	EXPONISTE	EXPONER	DESV1
PORTUGUÉS	5	PROPONAN	PROPONER	DESV4
PORTUGUÉS	6	PROPONÍ	PROPONER	DESV1
PORTUGUÉS	6	ESTOY	ESTAR	EST
PORTUGUÉS	6	ESTOY	ESTAR	EST
PORTUGUÉS	7	TIEM	TENER	DESV3
PORTUGUÉS	7	ENCUENTRAN	ENCONTRAR	EST
PORTUGUÉS	8	ES	SER	EST
PORTUGUÉS	8	ESTÁN	ESTAR	EST
PORTUGUÉS	8	ESTÁ	ESTAR	EST
PORTUGUÉS	8	HA	HABER	DESV3
PORTUGUÉS	8	SON	SER	EST
PORTUGUÉS	8	ESTÁN	ESTAR	EST
PORTUGUÉS	8	ESTÁ	ESTAR	EST
PORTUGUÉS	8	PARECEN	PARECER	EST
PORTUGUÉS	8	ESTÁN	ESTAR	EST
PORTUGUÉS	9	SON	SER	EST
PORTUGUÉS	9	DEMONSTRAN	DEMOSTRAR	DESV3
PORTUGUÉS	9	ES	SER	EST
PORTUGUÉS	10	ENCONTRASE	ENCONTRAR	DESV1
PORTUGUÉS	10	INCLUJENDO	INCLUIR	DESV3
PORTUGUÉS	10	ESTAN	ESTAR	EST

PORTUGUÉS	11	ENCONTRAN	ENCONTRAR	DESV1
PORTUGUÉS	11	ESTÁN	ESTAR	EST
PORTUGUÉS	11	ES	SER	EST
PORTUGUÉS	11	ES	SER	EST
PORTUGUÉS	12	ES	SER	EST
PORTUGUÉS	12	HAY	HABER	EST
PORTUGUÉS	12	ES	SER	EST
PORTUGUÉS	13	DEMONSTRÉ	DEMOSTRAR	DESV3
PORTUGUÉS	13	HAY	HABER	EST
PORTUGUÉS	13	ESTÁN	ESTAR	EST
PORTUGUÉS	13	DEBERÍ	DEBER	DESV4
PORTUGUÉS	14	ESTABA	ESTAR	EST
PORTUGUÉS	14	TIBE	TENER	DESV3
PORTUGUÉS	14	ES	SER	EST
PORTUGUÉS	14	ES	SER	EST
PORTUGUÉS	14	ANDAN	ANDAR	EST
PORTUGUÉS	14	ESTAE	ESTAR	DESV4
PORTUGUÉS	14	ESTOE	ESTAR	DESV4
PORTUGUÉS	14	ESTAE	ESTAR	DESV4
PORTUGUÉS	14	ESTOE	ESTAR	DESV4
PORTUGUÉS	14	ESTOE	ESTAR	DESV4
PORTUGUÉS	15	DESEMBARQUEI	DESEMBARCAR	DESV3
PORTUGUÉS	15	EMBARQUEI	EMBARCAR	DESV3
PORTUGUÉS	15	LLEGUEI	LLEGAR	DESV3
PORTUGUÉS	15	ENCUENTREI	ENCONTRAR	DESV3
PORTUGUÉS	15	ESTABA	ESTAR	EST
PORTUGUÉS	15	TENÍAMOS	TENER	EST

PORTUGUÉS	15	ENCONTRA	ENCONTRAR	DESV1
PORTUGUÉS	15	ENCUENTREI	ENCONTRAR	DESV3
PORTUGUÉS	15	ESTAREI	ESTAR	DESV3
PORTUGUÉS	15	ENCUENTRO	ENCONTRAR	EST
PORTUGUÉS	15	ES	SER	EST
PORTUGUÉS	16	ESTIBE	ESTAR	DESV3
PORTUGUÉS	16	SINTIÓ	SENTIR	EST
PORTUGUÉS	16	IBA	IR	EST
PORTUGUÉS	16	FUE	IR	EST
PORTUGUÉS	16	SENTÍ	SENTIR	EST
PORTUGUÉS	16	ERA	SER	EST
PORTUGUÉS	16	FUE	SER	EST
PORTUGUÉS	17	LLEGUÉ	LLEGAR	EST
PORTUGUÉS	17	PERCEBI	PERCIBIR	DESV3
PORTUGUÉS	17	HABÍA	HABER	EST
PORTUGUÉS	17	FUE	SER	EST
PORTUGUÉS	17	HUBO	HABER	EST
PORTUGUÉS	17	TENEMOS	TENER	EST
PORTUGUÉS	17	REENCONTRAMOS	REENCONTRARSE	EST
PORTUGUÉS	17	SENTIMOS	SENTIR	EST
PORTUGUÉS	17	PARÉCEME	PARECER	EST
PORTUGUÉS	18	LLEGE	LLEGAR	DESV1
PORTUGUÉS	18	SENTÍ	SENTIR	EST
PORTUGUÉS	18	ESTABA	ESTAR	EST
PORTUGUÉS	18	ERAN	SER	EST
PORTUGUÉS	18	HABÍA	HABER	EST
PORTUGUÉS	18	HABÍA	HABER	EST

PORTUGUÉS	18	SENTÍ	SENTIR	EST
PORTUGUÉS	18	SONRIEN	SONREÍR	EST
PORTUGUÉS	18	PARECÍAN	PARECER	EST
PORTUGUÉS	18	SENTÍ	SENTIR	EST
PORTUGUÉS	18	VI	VER	EST
PORTUGUÉS	18	PARECÍAN	PARECER	EST
PORTUGUÉS	18	PARECIEN	PARECER	DESV3
PORTUGUÉS	18	ESTÁ	ESTAR	EST
PORTUGUÉS	18	PARECE	PARECER	EST
PORTUGUÉS	18	ESTÁ	ESTAR	EST
PORTUGUÉS	18	ESTOY	ESTAR	EST
PORTUGUÉS	18	QUIERO	QUERER	EST
PORTUGUÉS	19	EMPEZATO	EMPEZAR	DESV3
PORTUGUÉS	19	LLEGUÉ	LLEGAR	EST
PORTUGUÉS	19	ESTABA	ESTAR	EST
PORTUGUÉS	19	ESTÁ	ESTAR	EST
PORTUGUÉS	19	ESTAMOS	ESTAR	EST
PORTUGUÉS	19	ESTAMOS	ESTAR	EST
PORTUGUÉS	19	ES	SER	EST
PORTUGUÉS	19	ESTAMOS	ESTAR	EST
PORTUGUÉS	20	ESTÁ	ESTAR	EST
PORTUGUÉS	20	ESTÁN	ESTAR	EST
PORTUGUÉS	20	ENCUENTRAN	ENCONTRAR	EST
PORTUGUÉS	20	ESTIVE	ESTAR	DESV3
PORTUGUÉS	20	TIEN	TENER	desv3
PORTUGUÉS	20	ESTÁ	ESTAR	EST
PORTUGUÉS	20	FUÉSEMOS	SER	EST

PORTUGUÉS	21	FUE	SER	EST
PORTUGUÉS	21	CONOCÍA	CONOCER	EST
PORTUGUÉS	21	FUE	SER	EST
PORTUGUÉS	21	FUE	SER	EST
PORTUGUÉS	21	HABÍAMOS	HABER	EST
PORTUGUÉS	21	ESTABA	ESTAR	EST
PORTUGUÉS	21	TUVO	TENER	EST
PORTUGUÉS	21	TENGO	TENER	EST
PORTUGUÉS	21	ESTOY	ESTAR	EST
PORTUGUÉS	21	TENGA	TENER	EST
PORTUGUÉS	21	ESTOY	ESTAR	EST
PORTUGUÉS	21	SUPONGO	SUPONER	EST
PORTUGUÉS	21	ESTARÉ	ESTAR	EST
PORTUGUÉS	22	ES	SER	EST
PORTUGUÉS	22	TENE	TENER	DESV1
PORTUGUÉS	22	TENE	TENER	DESV1
PORTUGUÉS	22	TENE	TENER	DESV1
PORTUGUÉS	22	TENE	TENER	DESV1
PORTUGUÉS	22	ESTOE	ESTAR	DESV4
PORTUGUÉS	22	CUOSTA	COSTAR	DESV3
PORTUGUÉS	23	TIENEN	TENER	EST
PORTUGUÉS	23	TIENE	TENER	EST
PORTUGUÉS	23	TIENE	TENER	EST
PORTUGUÉS	23	PENSO	PENSAR	DESV1
PORTUGUÉS	23	ES	SER	EST
PORTUGUÉS	23	TIENE	TENER	EST
PORTUGUÉS	24	ESTOY	ESTAR	EST

PORTUGUÉS	24	HAY	HABER	EST
PORTUGUÉS	24	ES	SER	EST
PORTUGUÉS	24	TIENE	TENER	EST
PORTUGUÉS	24	ES	SER	EST
PORTUGUÉS	24	SON	SER	EST
PORTUGUÉS	24	TIENE	TENER	EST
PORTUGUÉS	24	ESTÉN	ESTAR	EST
PORTUGUÉS	24	SOY	SER	EST
PORTUGUÉS	25	ÉS	SER	EST
PORTUGUÉS	25	ÉS	SER	EST
PORTUGUÉS	25	HAS SALIDO	SALIR	EST
PORTUGUÉS	25	ES	SER	EST
PORTUGUÉS	25	TENGO	TENER	EST
PORTUGUÉS	25	ES	SER	EST
PORTUGUÉS	26	ESTOY	ESTAR	EST
PORTUGUÉS	26	SABIA	SABER	EST
PORTUGUÉS	26	CONOCÍ	CONOCER	EST
PORTUGUÉS	26	FUE	IR	EST
PORTUGUÉS	26	HABÍA	HABER	EST
PORTUGUÉS	26	OFRECEN	OFRECER	EST
PORTUGUÉS	27	VINIERON	VENIR	EST
PORTUGUÉS	27	ERA	SER	EST
PORTUGUÉS	27	TENÍA	TENER	EST
PORTUGUÉS	27	ESTABA	ESTAR	EST
PORTUGUÉS	27	ERA	SER	EST
PORTUGUÉS	27	VE	VER	EST



## Agradecimientos

En primer lugar agradezco a Chichina por ser mucho más que mi directora de tesis, no sólo por iniciarme en la labor investigativa y contagiarme su pasión por la lingüística computacional si no por ser una guía y mi gran consejera.

En segundo lugar a todos los que me ayudaron de una u otra manera con la investigación, especialmente a Celina Beltrán, por colaborar con el análisis estadístico, a Rodolfo Bonino por compartir ideas y ofrecerme su ayuda desinteresada, a Barbi por acompañarme en tantas horas de estudio...

A mi familia y a mis amigas por brindarme su apoyo y darme aliento para seguir...

A mis amores Tati y Jose ...

A mi madre, que aunque no está conmigo me acompaña en cada paso que doy...

